AD_____
(Leave blank)

Award Number:
W81XWH-05-1-0298

TITLE:
Proteomic Prediction of Breast Cancer Risk: A Cohort Study

PRINCIPAL INVESTIGATOR:
Dr. Thomas E. Rohan

CONTRACTING ORGANIZATION:
Albert Einstein College of Medicine
Bronx, NY 10461

REPORT DATE:
March 2009

TYPE OF REPORT:
Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: (Check one)

X Approved for public release; distribution unlimited

☐ Distribution limited to U.S. Government agencies only; report contains proprietary information

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
28-03-2009	Final	1 MAR 2005 - 28 FEB 2009

4. TITLE AND SUBTITLE
Proteomic Prediction of Breast Cancer Risk: A Cohort Study

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W-81XWH-05-1-0298

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dr. Thomas E. Rohan

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Albert Einstein College of Medicine
Bronx, NY 10461

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
US Army Medical Research &
Materiel Command
Fort Detrick, MD 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release (Distribution Unlimited)

13. SUPPLEMENTARY NOTES

14. ABSTRACT.
Please see next page.

15. SUBJECT TERMS
Epidemiology, cohort study, proteomics

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code) 718-430-3355*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Abstract

Our objective is to develop and test proteomic methods for the prediction of breast cancer risk, an approach that has not been attempted previously. Our underlying hypothesis is that proteomic analysis of serum will identify proteins differentially expressed in women who do versus those who do not develop invasive breast cancer, and that these differences will be identifiable prior to the clinical presentation of breast cancer. Our work is being conducted in two phases, a training phase and a test phase. Both phases will be conducted as case-control studies nested in a population-based cohort of women who were members of Kaiser Permanente. These serum specimens were collected between 1986 and 1992. We have finalized our cohort definition, finalized the definition of cases and controls; finalized the criteria for matching controls to cases; selected the cases and controls; pulled and aliquotted the serum specimens. For the proteomic analysis, we have developed a detailed protocol for analysis of the serum samples. Briefly, the serum sample is loaded onto an immunoaffinity column to deplete twelve abundant proteins, and the flow-through fraction is collected and subjected to tryptic digestion. Subsequently, the peptides are labeled with iTRAQ reagents and fractionated by cation exchange chromatography (SCX). Six pooled SCX fractions are separately loaded onto a reverse phase column and followed by MALDI-TOF/TOF (4800 Proteomic Analyzer) analyses. The data collected are automatically processed, combined, and searched against a human protein database. This procedure has been thoroughly tested for quantitation and complexity (dynamic range) and a reproducibility study is underway. By applying this high-resolution proteomic approach to a prospective setting, this ongoing project should enhance our ability to identify those women at increased risk of breast cancer and intervene before they progress to cancer. Furthermore, it is expected to provide insight into the biological processes underlying breast cancer development through the identification of protein markers of disease and disease susceptibility genes.

Table of Contents

Introduction

New approaches to breast cancer risk prediction are needed given the modest discriminatory accuracy of existing risk prediction models at the individual level. Our objective is to develop and test proteomic methods for the prediction of breast cancer risk, an approach that has not been attempted previously. Our underlying hypothesis is that proteomic analysis of serum will identify proteins differentially expressed in women who do versus those who do not develop invasive breast cancer, and that these differences will be identifiable prior to the clinical presentation of breast cancer. Specifically: (1) using a training set of serum specimens sampled from a population-based cohort of women who were members of Kaiser Permanente, who had their blood samples taken at a multiphasic health examination (MHC) between 1986 and 1992**,** and who were followed up to determine subsequent breast cancer occurrence, we will apply highly sensitive proteomic approaches in order to identify biomarkers that discriminate between women who developed invasive breast cancer within 5 years of having a serum sample collected (cases) and women who remained free of breast cancer for at least that long (controls); (2) using a validation set of serum samples collected from a separate group of cases and controls in the cohort, we plan to test the biomarkers identified using the training set. Nested case-control studies will be undertaken at the training (32 cases/32 controls) and validation stages (20 cases/20 controls). Cases will be women with no history of breast cancer at baseline who developed a subsequent incident, invasive breast cancer within 5 years of collection of a serum sample; controls (no breast cancer history) will be selected from the same cohort using risk-set sampling and matched individually to cases on calendar year of the MHC exam at which the serum sample of interest was collected, age at that MHC exam, and time since last meal. To reduce the possible impact of breast cancer heterogeneity on interpretation of the results, we have restricted inclusion to white, postmenopausal women; should the results prove informative, we plan to study other subgroups (e.g., African American, premenopausal, etc.) in the future. Data on breast cancer risk factors will come from questionnaire and physical exam data gathered at the MHC. All proteomic spectra will be analyzed by the Mascot program to identify the protein sequences generating all peptide ions. After biomarker ions have been identified, validation will be carried out by the same analytical procedures following only the biomarker ions of interest. The laboratory work will be performed “blinded” to case-control status. Supervised clustering algorithms such as support vector machines will be used for the statistical analysis of the training dataset to identify candidate marker patterns that best discriminate between cases and controls. The most promising markers identified in the training set will be applied to subjects in the validation set and used to classify them as cases, controls, or neither. In contrast to the training set, where case-control status will be known to the data analysts, case-control status will not be known during analysis of the validation set. Therefore, we will determine how well the method developed at the training stage discriminates between cases and controls in the validation set.

Body

(i) Eligibility criteria

The cohort consists of white, postmenopausal women aged 55 to 80 years at the time of a blood draw at Kaiser Permanente between 1986 and 1992. None of the samples from the cohort had been thawed prior to use by us. As indicated in our original proposal, we chose to restrict attention to white women to reduce the possible impact of breast cancer heterogeneity on interpretation of the results.

(ii) Case/control definition and selection

The work for this project is being conducted as nested case-control studies, both at the training and at the validation stages. Cases are defined as white, postmenopausal women with no history of breast cancer at the time of recruitment. Inclusion is restricted to subjects between the ages of 55 and 80 years. We identified 68 potentially eligible cases by merging data from the multiphasic cohort/serum repository databases with data from the Kaiser Permanente tumor registry. The 52 cases selected for the study are a random sample of all eligible cases. Controls are matched 1:1 to the corresponding case. They are white, postmenopausal women with no history of breast cancer and who have not developed breast cancer by the date of diagnosis of the corresponding case. The controls were selected using risk-set sampling with replacement. They were matched to the corresponding case on age (within 1 year), date of serum collection (to within 1 month), and time since last meal (0-3 hours, 4-9 hours). Furthermore, cases and controls were matched with respect to membership of Kaiser Permanente, in the sense that as with the cases, controls were required to have been members from one year prior to serum collection and to have been a member at the time of diagnosis of the corresponding case (those whose membership lapsed for a period exceeding 3 months ceased to be eligible to be selected as a control).

(iii) Data file

Study identification numbers were assigned to the cases and controls. These numbers are linkable to the Kaiser IDs, which allowed extraction of corresponding covariate information from the cohort database. The data file containing the covariate information required for the analysis of the training set data has been created.

(iv) Aliquoting of serum samples

The pulling, testing (for dessication), and aliquoting of the serum specimens was completed at the Orentreich Foundation. For each subject included in the study, 5 x 20 μl aliquots were made. Furthermore, a common serum pool was created by adding 50 μl from each study subject to the pool. The pool serves as a common standard that is included in each run. For each case-control pair, 5 sets of samples were prepared, each set consisting of 20 μl aliquots for the case, the matched control, and the common pool. The location of the case and control aliquot within each triplet was assigned randomly and the laboratory staff are blinded to the identity of the case and control samples. The specimens were shipped to and received at the Albert Einstein College of Medicine where they are held in storage at -80^{o}C until they are needed for analysis.

(v) Development of methods for proteomic analysis of serum samples

Despite several technical difficulties, we have now established detailed protocols and reaction conditions for each step of the proteomic analysis. Specifically, we have developed an integrated, highly sensitive, protocol for proteomic analysis involving fractionation by immunodepletion, multi-dimensional HPLC, and analysis by MALDI-TOF/TOF-MS, to tackle the complexity and dynamic range of the serum proteome.

In the previous report, the depletion, labeling, and digestion steps had been worked out. In our preliminary data analysis, major abundant proteins were still found among the top ranks of the protein identity list. This is a common problem and is still documented in recent literature. To minimize the adverse effect of the major undesirable proteins, we have recently decided to change the single column (12-protein) depletion method to a dual column (14-protein column and SuperMix column, Genaway) approach. Other fractionation steps such as the strong cation exchange (SCX) and reversed phase-MALDI target spotting routines were also optimized. In the case of SCX, a more detailed procedure is adopted to overcome the gradient inconsistency problem that causes run-to-run variation. For reversed phase liquid chromatography (HPLC), an updated solvent gradient was used to ensure that peptides are eluted evenly in amount across the whole gradient. This will eventually reduce the ion suppression effect during MALDI spectra acquisition.

Another major update is that we decided to collect the SuperMix eluent from the dual column depleteion run since proteins of interest might be in it. According to the manufacturer, the SuperMix could remove 77 medium abundant proteins from serum.

The sequence of steps in this analysis is summarized in Figure 1, the detailed description of the procedure is provided in the Proteomic Laboratory Operation Manual (Appendix 1) and the checklists for each step (Appendix 2). Briefly, each serum sample is loaded onto an immunoaffinity column to deplete fourteen abundant and 77 moderately abundant proteins, and the flow-through fraction as well as the eluent of the SuperMix column are collected and subjected to tryptic digestion. Subsequently, the peptide digest is labeled with iTRAQ reagent, mixed with two other digests (to combine differentially labeled case, control, and pool), and subjected to strong cation exchange (SCX) chromatography. Each SCX fraction is then loaded onto the reverse phase column and pooled fractions spotted onto 384 well plates for MALDI-TOF/TOF (4800 Proteomic Analyzer) analyses. The data collected are automatically processed, combined, and searched against human protein databases. The raw spectral files are processed to provide appropriate bioinformatic data. All stages of this procedure have been developed and refined over the last year as described below.

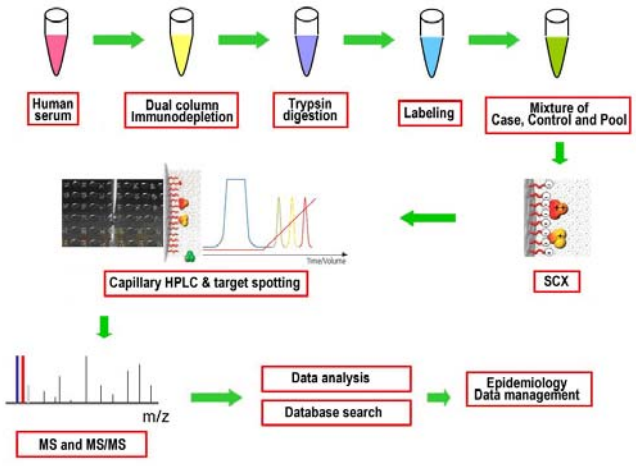


Figure 1. Steps in the proteomic analysis of breast cancer serum samples

(a) Sample preparation. To maximally remove the most abundant proteins from serum, we recently adopted the dual column approach. The serum sample (20 µl) is loaded onto the two linearly assembled immunodepletion columns (IgY14LC5 and SuperMix columns, Genway Biotech Inc, San Diego, CA), the flow-through fraction is collected (Fig. 2), and its volume is reduced by ultrafiltration using a spin filtration tube with a 5 kDa Mr cutoff. As compared to the depletion profile in Fig.3 which was obtained using the previous single column method, the newly adopted dual column approach has a higher efficiency in removing the abundant proteins (95% vs. 80%) The efficiency of the filtration is shown in Fig. 4. The concentrated flow-through fraction is digested overnight with 5 µg of trypsin in 20% acetrontile and 30 mM TEAB. Figure 4 shows the results of digestion performed under various conditions. The digest was subjected to labeling with one vial of iTRAQ reagent (114, 115 or 116) overnight at 37°C. Residual reagent is quenched by adding 420 µl of water and allowing excess reagent to completely hydrolyze over an additional 30 minutes. Afterwards, 20% of the labeled digest is saved for glycoprotein determination at -80°C. The three differently labeled samples are subsequently mixed. To avoid any incomplete labeling of peptides due to a possible bad batch of iTRAQ reagent, a brief test is performed by labeling a standard peptide, des-Arg-Bradykinin, prior to labeling the samples.

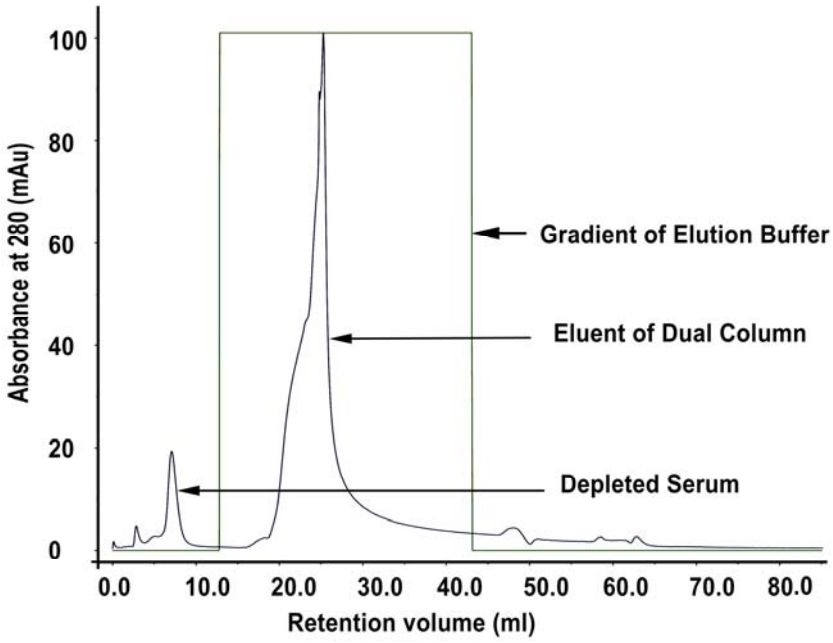


Figure 2. Immunoaffinity depletion of abundant proteins from serum using the newly adopted Dual Column approach. With this approach, about 95% of all proteins are removed after the run, indicating a higher efficiency in serum protein depletion.

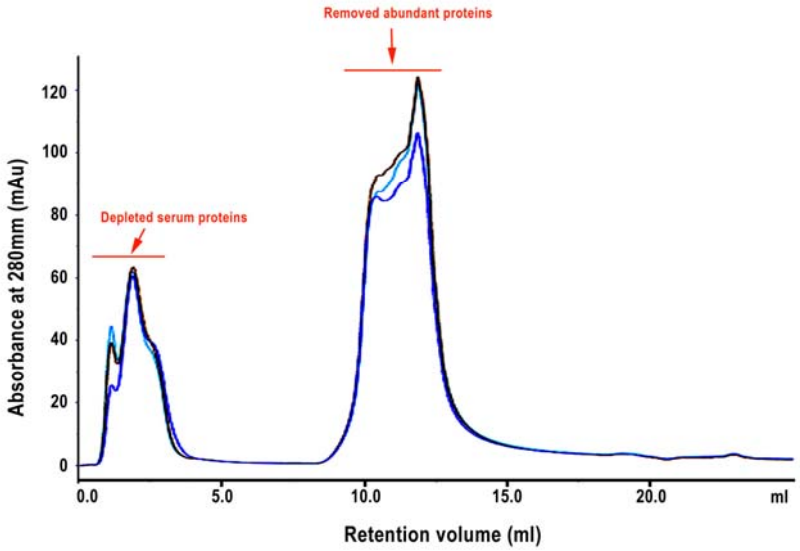


Figure 3. Immunoaffinity depletion of abundant proteins from serum using the single IgY12 protein depletion column. The overlay of three serum samples from one stratum is shown. The first eluting peaks are combined as the flow-through fraction for further analysis. In this combined fraction the majority of abundant proteins are removed.

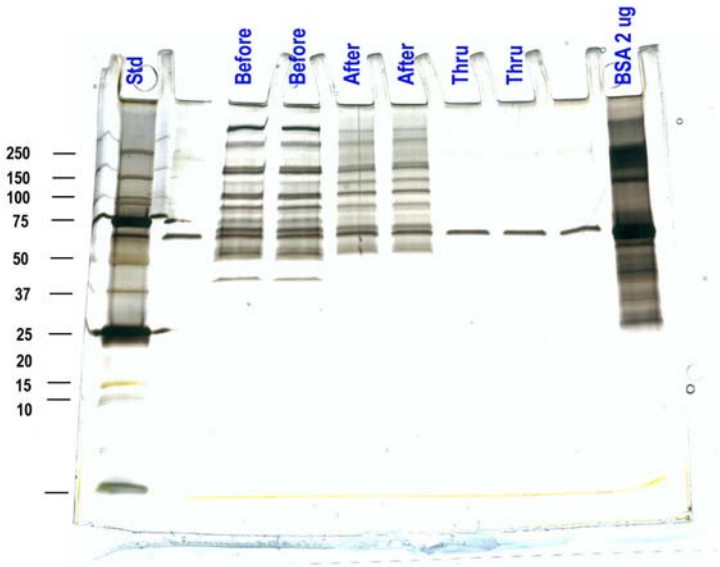


Figure 4. SDS-PAGE profile of depleted proteins before (Before) and after (After) using a filtration-concentration spin tube with a 5-kDa Mr cutoff. Equivalent amounts of samples were loaded into each lane of the gel in duplicate. The gel was silver stained revealing the resolved proteins. The spin flow-through (Thru) of the sample is also shown.

Several changes to the original procedure were necessary for optimization, including the elimination of the reduction and alkylation step of the immunodepleted flow-through fraction under denaturing conditions and its subsequent concentration on a C4 column (complete removal of guanidium hydrochloride was difficult and adversely affected the C4 column step). The end result was to streamline the procedure. The optimization of the biochemical steps was completed using standard proteins and the NIST serum standard, often used for proteomics methods development. The reduction in the amount of sample that is labeled, significantly improved labeling efficiency (to > 95%). Only 40% of the sample is now labeled with the remaining 60% being saved. A 40% proportion, corresponding to 33% of the total, proceeds to the 2D LC step. A proportion corresponding to 7% is saved for glycoprotein biomarker analysis using multi-lectin affinity chromatography. The obtained glycopeptides will be treated with PNGase-F to identify the peptides that possessed the N-linked sugars (Yang Z, Harris LE, Palmer-Toy DE, Hancock WS. Multilectin affinity chromatography for characterization of multiple glycoprotein biomarker candidates in serum from breast cancer patients. Clin Chem. 2006;52:1897-905. Plavina T, Wakshull E, Hancock WS, Hincapie M. Combination of abundant protein depletion and multi-lectin affinity chromatography (M-LAC) for plasma protein biomarker discovery. J Proteome Res. 2007;6:662-71). This glycoprotein analysis was not part of the original proposal, but may provide additional useful information. As can be seen in the detailed protocol (Appendix 1), other modifications to the sample preparation protocol were also made.

(b) Off-line 2D LC coupled with MALDI-TOF/TOF analyses. In the first dimension, the combined peptide mixture is separated by strong cation exchange (SCX) chromatography on an AKTA Purifier 10 system (GE Healthcare Bio-Sciences, Piscataway, NJ) using a PolySulfoethyl A ^{TM} column (2.1 × 100 mm, 5 μm, 300Å; Poly LC Inc, Columbia, MD). The sample is diluted in 4.3 ml of SCX loading buffer (20% acetonitrile, 10 mM potassium phosphate, pH 3) and loaded onto the column. The column is then washed isocratically for 20 min at 0.1 ml/min to remove excess reagent. Peptides are eluted with a multistep gradient of 0-700 mM KCl (in 20% acetonitrile, 10 mM potassium phosphate, pH 3) over 15 minutes at a flow rate of 0.1 ml/minute, with fractions collected at 1-minute intervals (Fig. 5). The second dimension of the peptide separation is performed on an Ultimate^{TM} 3000 chromatography system equipped with a Bai Probot MALDI spotting device (Dionex, Sunnyvale, CA). Six individual SCX fraction pools are injected and

captured onto a trap column (1× 15 mm; Dionex, Sunnyvale, CA) and then eluted onto an RP –C_{18} capillary column (300 µm ×150 mm; Dionex, Sunnyvale, CA) with a gradient of buffer B (buffer A, 0.1% TFA, 5% acetronitrile, 95% H_2O; buffer B, 0.1% TFA, 95% acetonitrile, 5% H_2O). Column effluent is mixed automatically in a 1:1 ratio with premade MALDI matrix, 6.2 mg/ml of alpha-cyano-4-hydroxycinnamic acid in 36% methanol, 56% acetonitrile, and 8% H2O (Agilent, New Castle, DE) using a probot MALDI spotting device and spotting directly onto MALDI plates. MALDI plates are analyzed on an AB 4800 Proteomic Analyzer from Applied Biosystems (Framingham, MA).

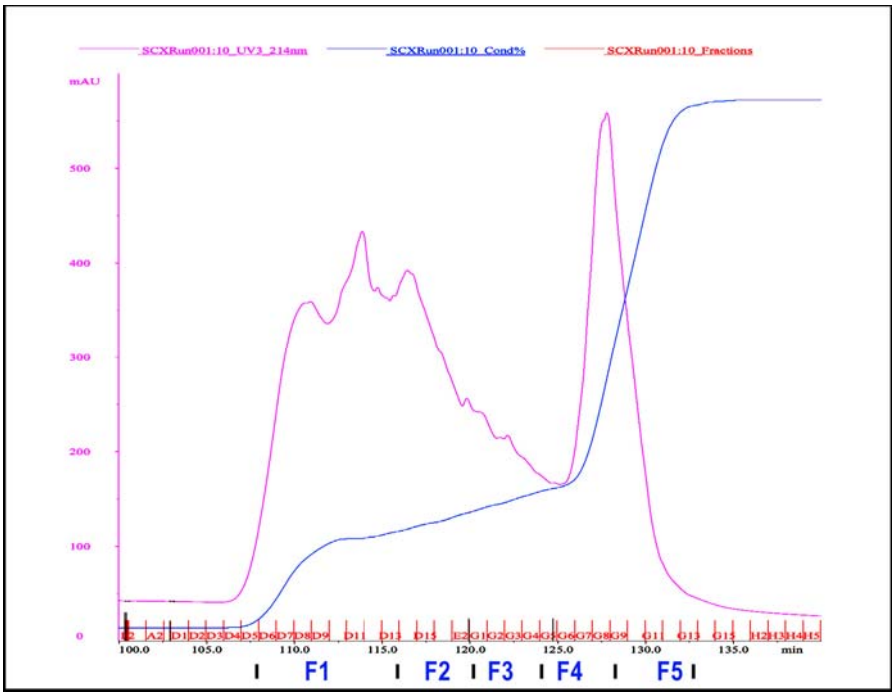


Figure 5. SCX chromatogram of iTRAQ-labeled digest of Stratum 01, showing absorbance at 214 nm (purple) and the KCl gradient (blue, 0 - 700mM) and the 5 fraction pools subjected to capillary LC and spotting onto 5 TOF/TOF target plates

In the development of this step, the slope of the gradient was optimized so that an approximate equivalent amount of peptide ions would be detected in each of the fraction pools. This process entailed testing of various gradients by capillary HPLC and spotting the eluent onto TOF/TOF target plates, followed by data acquisition and comparison of the data sets. Each SCX fraction is deposited onto one target of 384 fractions. Optimization of the SCX chromatography condition was followed by optimization of the capillary HPLC gradient to distribute the peptides as evenly as possible across the TOF/TOF target plates.

Since with the newly upgraded 4800 system we can use the 384 well MALDI plate, the LC run method is also optimized to allow maximal sample collection with improved peptide separation. The LC run flow rate is changed from 2 µl/min to 4 µl/min, collection time is changed from 40-90 min duration to 35-90 min duration. The effect of this flow rate change on the peptide elution profile in shown in Fig. 6.

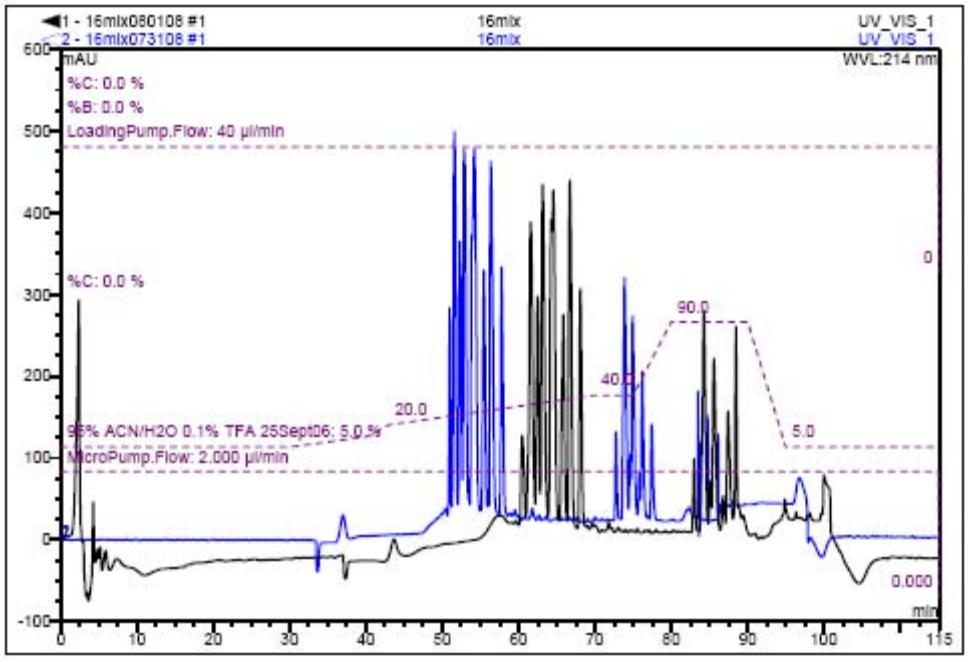


Fig. 6 Comparison of LC profile of a standard peptide mixture run at 4 µl/min with that at 2 µl/min. A faster flow rate leads to earlier elution of peptides. Blue curve, run at 4 µl/min; Black curve, run at 2 µl/min.

(c) Data processing. Data analysis is performed using in-house software (Du P, Angeletti RH. Automatic deconvolution of isotope-resolved mass spectra using variable selection and quantized peptide mass distribution. Anal Chem., 2006, 78:3385-92, 2006, P Du, R Sudha, MB Prystowsky, R Hogue Angeletti; Data reduction of isotope-resolved LC-MS spectra. Bioinformatics, 2007, Jun 1;23(11):1394-400) to build arrays with SCX fraction number, retention time, and precursor monoisotopic mass of all ions containing the iTRAQ reporter ions (114, 115 and 116), along with their signal intensities. The retention time drift is corrected for by using the retention time of the internal standard peptides. Peptides having similar mass (± 0.1 Da) and retention time (± 60 sec) are considered to be the same peptide. Intensity levels of ions at m/z 117 are used as indicators of noise level. In analyses of replicate samples of the pool, a high level of reproducibility was observed for all iTRAQ reporter ions having intensities greater than 500 counts (Fig. 7).

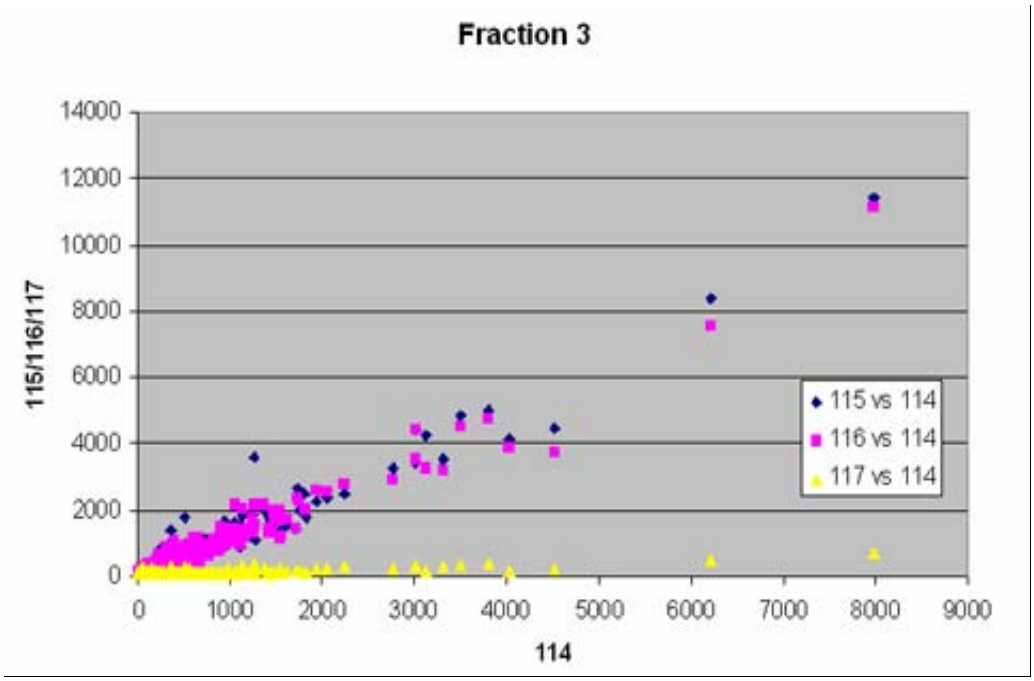


Figure 7. Plot of the iTRAQ reporter ion intensities (*m/z* 114, 115 & 116) of replicate pooled serum samples. The ion intensity value plot of *m/z* 117 depicts our baseline noise level. Ideally, the ion intensity values from *m/z* 114, 115 & 116 should all be equal to each other. The calculated correlation coefficients are: 114:115, 0.9701; 114:116, 0.9719; and 115:116, 0.9811.

In addition, the data are searched through the TOF/TOF GPS server and Mascot to identify the sequenced peptides and their parent proteins. As an example, from the stratum 02 immunodepleted serum, more than 1,098 proteins can be identified with a Confidence Interval (C.I.%) of 95% or greater. It should be noted that in the quantitative analysis of serum, most peptides will be unchanged. In the present experiments, this will mean that in most cases the 114 and 115 marker ions (representing blinded case and case control samples) should be indistinguishable from the 116 marker ion, the pooled control (Figs. 8 and 9). Of the 1,098 proteins identified, 305 proteins are highly relevant to pathways and disease processes and are not in the upper tier of abundant proteins (Appendix 3).

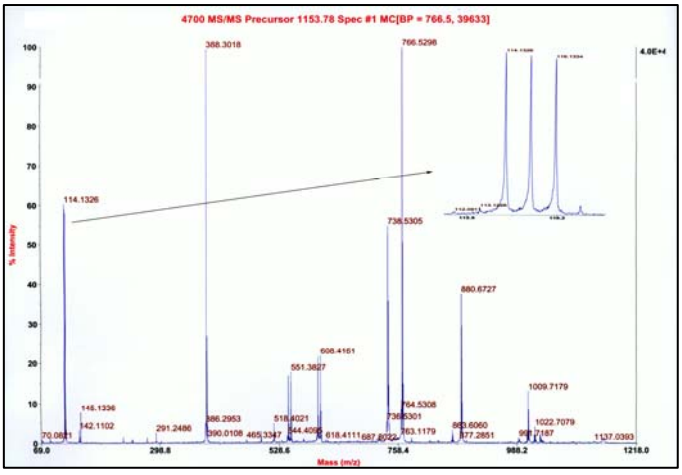


Figure 8. MS/MS spectrum of a peptide ion with approximately equal intensities of the iTRAQ reporter marker ions for case, control, and pooled sample. The inset is a magnification of these marker ions (*m/z* 114, 115 & 116)

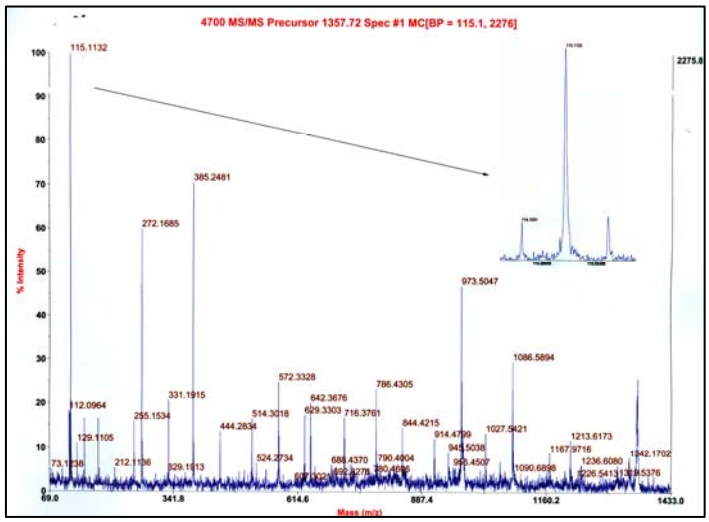


Figure 9. MS/MS spectrum of a peptide ion with an intense 115 iTRAQ reporter marker ion (inset), indicating the increased expression of this peptide in either the case or control serum.

Using the Protein Interrogation of Gene Ontology and KEGG databases web-based program (PIGOK; http://128.40.158.133/pigok.html) the proteins identified from part of stratum07 (Fractions 4 and 5, 500 hit search, minimum modifications) were categorized and are shown below in figures 10 to 12.

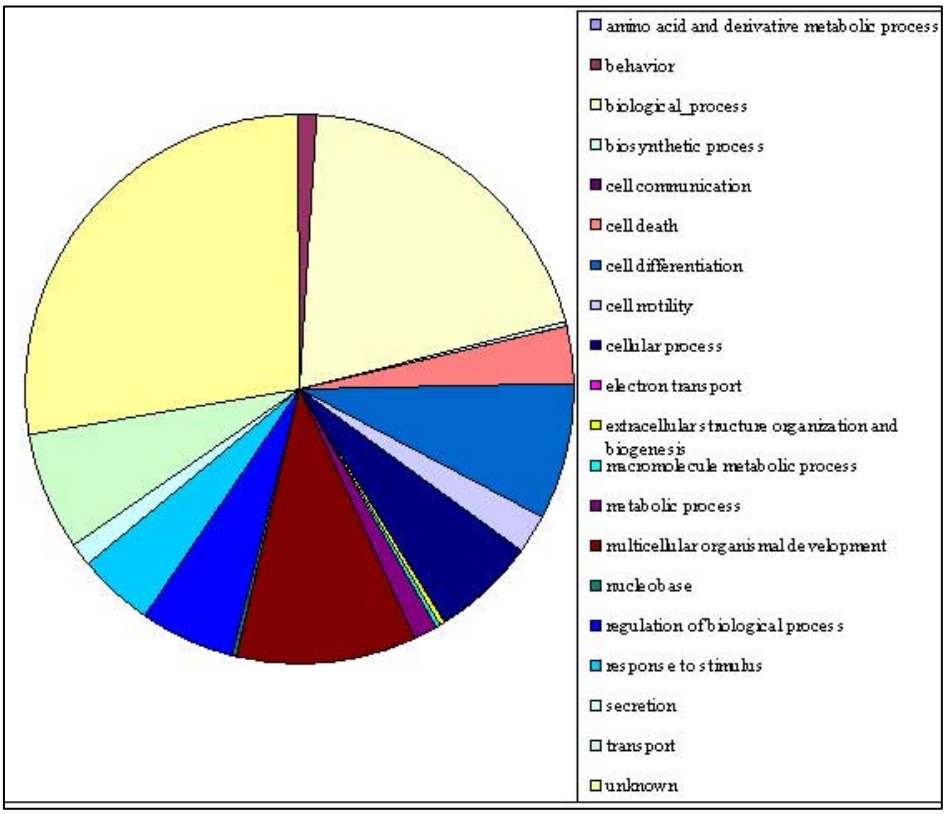


Figure 10. Categorization of proteins according to their molecular functions using PIGOK. These proteins are identified using the GPS Explorer program. MSMS data were taken from SCX fractions 4 and 5 of stratum7.

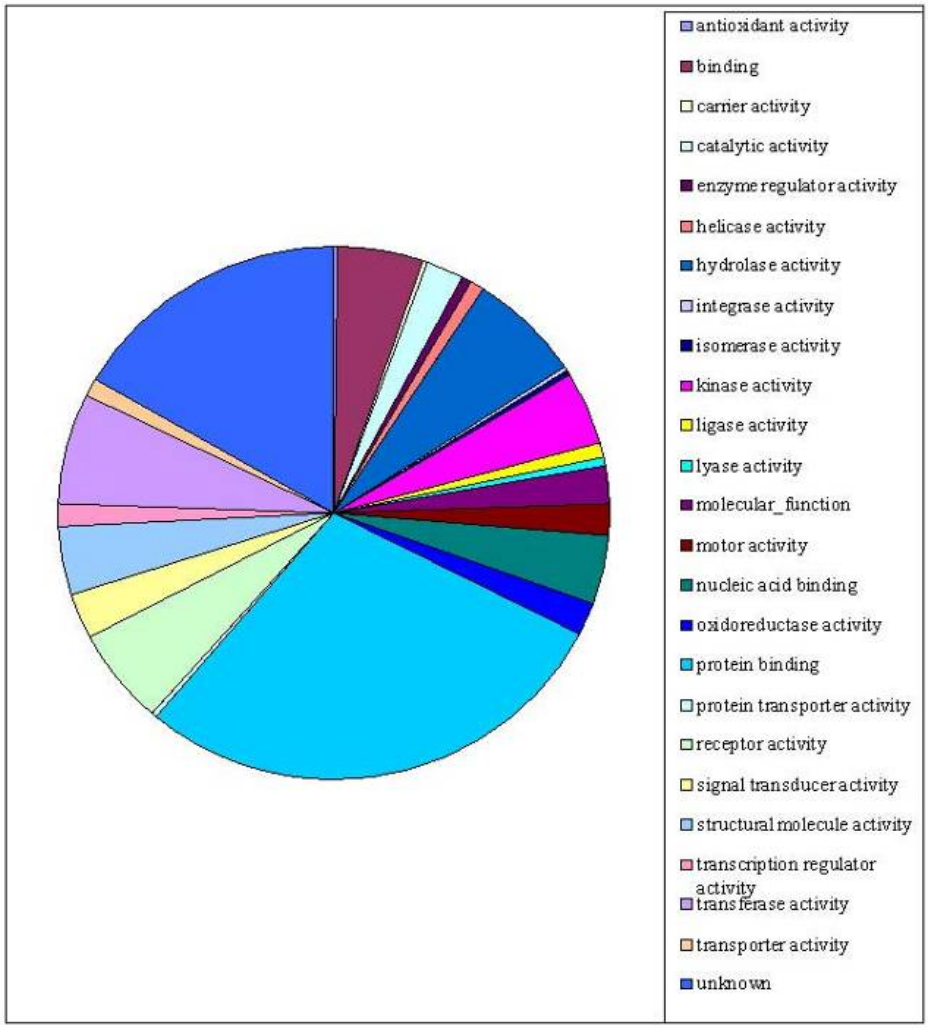


Figure 11. Categorization of proteins according to their biological process using PIGOK. These proteins are identified using the GPS Explorer program. MSMS data were taken from SCX fractions 4 and 5 of stratum7.

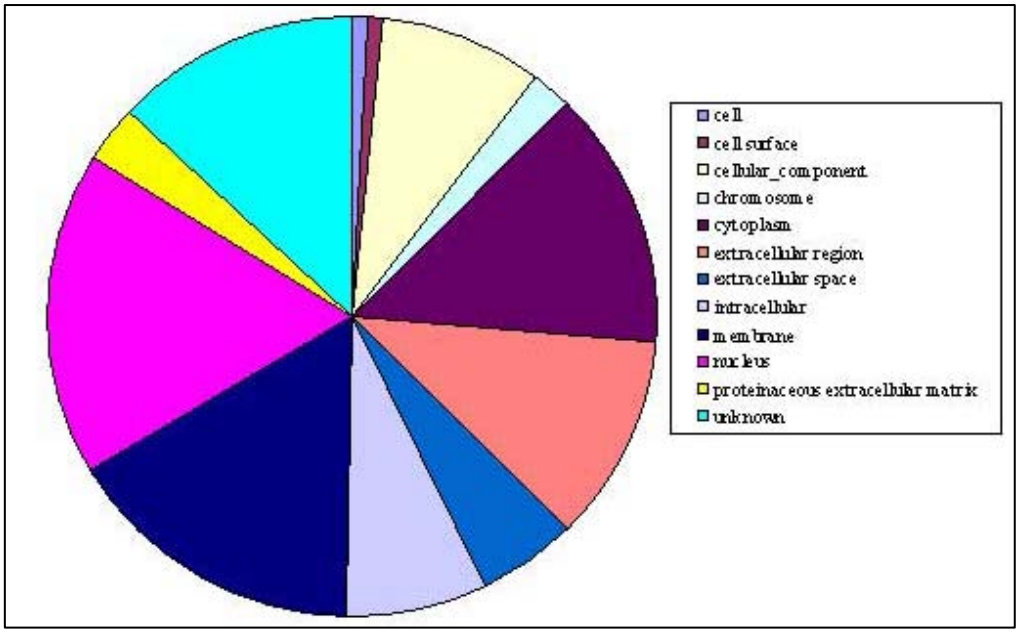


Figure 12. Categorization of proteins according to their cellular_component using PIGOK. These proteins are identified using the GPS Explorer program. MSMS data were taken from SCX fraction s 4 and 5 of stratum7.

It is notable that low abundance proteins are being consistently identified in these experiments. For example, proteins such as NF-IL6 and interferon $\alpha 2/\beta 1$ are reported to be present at levels less than 200 pg/ml, whereas serum albumin is present at approximately 35-50 mg/ml. This would represent a dynamic range of detection of almost 10^8 orders of magnitude.

(vi) Data and operations management

The process of tracking samples and data throughout the complicated process from the initial immunodepletion phase through the MSMS phase is crucial in this study. Each of the procedures in this process has been documented in an operations manual and checklists have been devised for each step of the process (Appendices 1 and 2) to ensure that all procedures are in place and are adhered to strictly. In order to track the many different test tubes that are produced at each juncture, cryo-labels have been created that clearly identify the current procedure and the sample ID number.

Data are stored on three different computers, each of which is connected to the equipment necessary for each step of the process. Directories have been created in each of these machines to clearly identify project data and these directories are backed up weekly onto a portable hard drive. After analysis of each case-control stratum is completed, the hard drive is copied to the data management servers and data on the hard drive are deleted so that data resulting from the analysis of the next stratum can be backed up.

During the MSMS phase, data files for each of the six fractions are created. These files are then converted into separate Td2 files in order to prepare analysis files. An analysis program is run on each of these files and space delimited text files are outputted which contain one line of data for each spectrum in the given fraction. Data included in this file include the spot number (spectrum), M/Z, mascot score, protein accession number, and the intensity for each sample by peptide. Quality assurance procedures include steps to ensure that the file is created correctly before raw data files are moved. After these interim analysis files are created, all data for each stratum are merged into one file which will be used to create the final analysis file for the study.

Project management is facilitated by a portal web site which has been created for this study and allows all members of the study to collaborate in a shared environment. The portal site hosts a project calendar and a tracking system which allows each member to update the status of each sample as it is completed and document issues that occur. The data manager is automatically notified when this system is updated to ensure that the project is tracked efficiently and monitored appropriately.

Key research accomplishments

We have developed an integrated, highly sensitive, proteomic analysis involving fractionation by immunodepletion and multi-dimensional HPLC, analysis by MALDI-TOF/TOF-MS and sophisticated computational analysis of the data, to tackle the complexity and dynamic range of the serum proteome. We have shown that this procedure identifies proteins that are highly relevant to pathways and disease processes and that are not in the upper tier of abundant proteins. We have established a sophisticated procedure for operations management and quality control.

Reportable Outcomes

So far our publications have been on the development of algorithms for signal processing:
1. Du P, Angeletti RH. Automatic deconvolution of isotope-resolved mass spectra using variable selection and quantized peptide mass distribution. Anal Chem., 78:3385-92, 2006,
2. P Du, R Sudha, MB Prystowsky, R Hogue Angeletti (2007) Data reduction of isotope resolved LC-MS spectra. Bioinformatics, in press)

In addition, we presented an abstract describing our work at the DOD Era of Hope meeting in Baltimore in June, 2008: Xiao Y, Nieves E, Yeung Y-G, Du P, Ginsberg MS, Habel L, Vogelman J, Orentreich N, Angeletti RH, Stanley ER, Rohan TE. Serologic biomarker discovery for human breast cancer using a proteomics approach – method development. Department of Defense, Breast Cancer Research Program, Era of Hope Meeting, Baltimore, June 2008.

Finally, a manuscript describing our proteomic method is in preparation (Appendix 4): Y. Xiao et al. Serologic biomarker discovery for human breast cancer using proteomics approach-method development.

Personnel receiving pay

The following personnel received pay from the research effort for at least part of the project duration: Thomas Rohan, Ruth Angeletti, Peicheng Du, Charles Hall, Nika Heinz, Lianji Jin, Liyuan Ma, E. Richard Stanley, Y.G.Yeung, Yansen Xiao.

Conclusions

We completed the MS analysis of the first 9 strata using an earlier version of our protocol. However, this revealed some limitations of the approach and led to the revised protocol described above. A reproducibility study based on the revised protocol is currently ongoing on 6 of the pooled samples, with 3 in each of 2 runs, to mimic 2 strata (i.e., 3 pooled samples in one stratum and then 3 pooled samples in another). This will yield measures of intra- and inter-run repeatability. Our need for intense data acquisition to complete this project revealed shortcomings of the mass spectrometer (Applied Biosystems 4700 Proteomics Analyzer) used for initial data collection, namely its inability to operate full-time due to charging effects now known to affect instrument performance. This limited the number of plates that were able to be used per day and per week before either needing to allow the instrument to discharge or having to break the vacuum and clean. Because of these problems, which delayed the project, we traded-in the 4700 instrument for the 4800 model, which we have determined is able to operate full-time, but will still require regular cleaning. While we did fall behind schedule, we believe that we have developed procedures that provide the complexity required to uncover new markers of disease and allow analysis of the actual study specimens at a reasonable rate. The personnel involved are highly trained and familiar with all aspects of the project. Our goal now is to complete our repeatability study, finalize and publish the manuscript describing our proteomic methods, and to seek funds from other sources to allow us to complete the project.

Appendix 1

Proteomic Lab Operations Manual

Table of contents

I. Test tube preparation

A. For the first phase of this study, we will be analyzing 40 case control pairs. Each case control pair has been assigned a stratum number (01-40). Each sample has a unique ID number (1-80). The lab will receive 5 aliquots for each one of the samples. For each case-control pair (i.e. stratum), the lab will also receive a pooled sample. The labeling schema developed for this study identifies each test tube as: STRATUM# _ ID # _ ALIQUOT #

Example: For the first stratum (01), the lab will receive 3 test tubes.

- The first test tube will be labeled 01_1_1. 01 represents the stratum number. The 1 in the middle represents the ID for this sample. The last 1 represents the 1st aliquot for this sample.
- The second test tube will be labeled 01_2_1. 01 represents the stratum number. The 2 represents the ID for this sample. The last 1 represents the 1st aliquot for this sample.
- The third test tube will be the pooled sample for the 01 stratum and will be labeled 01_PL_1. 01 represents the stratum number. The PL represents POOL for this sample. The last 1 represents the 1st aliquot for this pooled sample.

B. A label sheet will be prepared by the Department of Epidemiology and Population Health (DEPH) for each stratum. The sheet will contain all of the labels that are needed for the process. Each label contains the procedure step name (CONC, DIGEST, iTRAQ etc) and the numbers used to identify the sample. The first 2 columns of the label sheet contain the labels for the <u>first</u> aliquot of each stratum. The other 4 columns contain extra labels for the other aliquots in case they are needed. As a rule, only the first 2 columns of the label sheet will be used unless a 2nd aliquot is needed for a particular sample.

II. Immunodepletion
(Removal of major abundant serum proteins by immunodepletion chromatography, 3 serum samples consecutively at a time)

Buffers

Buffer A1: Phosphate Buffered Saline (PBS), pH 7.4, from Sigma (P3813-10PAK, contains 0.01 M phosphate, 0.138 M NaCl, 0.0027M KCl). Freshly prepare before use for the day.

Buffer B1: Stripping Buffer: 0.1 M glycine-HCl, pH 2.5, containing 0.02% sodium azide.

Neutralization Buffer A2: 0.1M Na_2HPO_4-NaH_2PO_4, pH 8.0, containing 0.02% sodium azide.

Column Storage Buffer: PBS containing 0.02% sodium azide.

Filter all buffers through a 0.2 µm sterile membrane filter and store at 4°C, except buffer A1 which need to be prepared freshly each time.

Columns

1. Column One: GenWay Seppro IgY 14 LC5 affinity column (Cat.# 28-288-12014-LC5), 12.7 x 39.5 mm. Removes human Albumin, IgG, a1-antitrypsin, IgA, IgM, Transferrin, Haptoglobin, a1-Acid glycoprotein (Orosomucoid), a2 Macroglobulin, HDL (Applipoprotein A-I & A-II), Fibronogen, Complement C3 and LDL (mainly Apolipoprotein B). Capacity: 100 – 125 ul human plasma or serum per injection.

2. **Column Two**: GenWay SuperMix LC1 affinity column (Cat.# 28-288-23078-LC1), 6.4 x 31.5 mm. Target to remove at least 77 moderately abundant proteins from human plasma. Capacity: 100-125 ul human plasma or serum per injection.

3. **The columns must be stored at 4°C in column storage buffer. The columns must be washed with 3 column volumes of room temperature degassed dilution buffer IMMEDIATELY after it is taken from 4°C storage**. This procedure will prevent air bubble formation inside the column.

Instrumentation

AKTA purifier
Injection loop size: 250 µl
Flow cell (Cat. 18-1147-25): 3 mm pathlength, 0.7 µl hold up volume.
Mixer size: 0.6 ml
(Check the synchronization of UV tracing and sample collection of the fraction collector)
Syringe: 250 µl syringe from Hamilton

Procedure

1. Bring PBS, Stripping buffer, neutralization buffer and column storage buffer to room temperature and re-dissolve precipitates. (Phosphate precipitates when stored at 4°C). Degas all buffers by purging the buffers with helium for 10 minutes.
2. Turn on AKTA work-station and computer. Wait for system synchronization to complete.
3. Add 4 ml of 5% Triton X-100 to each inner tube of 12 Ultra-4 centrifugal concentrators (Millipore, Bedford, MA 01730. Size: 5,000 MWCO. Cat. No. UFC800596, 96pk) and leave them on ice for >30 min. These concentrators will be used in Step 30 in the Trypsin Digestion section.

4. Connect the 250 µl sample loop to the correct valve inlet position.
5. Connect UV light, pH meter, and Conductivity meter to the correct position.
6. Run the *equilibrate-system* program under Method, to bring all solutions to running-condition.
7. Switch valve to "inject" position to wash and purge the loop with 2 ml of Buffer A1. Then switch valve to "load" position to wash and purge the loading syringe adaptor and the loop with 2 ml of Buffer A1 from a syringe fitted with the loading needle.
8. Connect the LC5 column immediately after brought out from 4°C, and purge with 20 ml of Buffer A1 at 2ml/min. Make sure that no air is introduced to the column during connection.
9. Then connect the LC1 column in serial with LC5 immediately after brought out from 4°C, and purge with 5 ml of Buffer A1 at 2 ml/min. Make sure that no air is introduced to the column during connection.
10. Test the fraction collector tubing to make sure it's not clogged by switching the Frac to F2 and allowing 5 ml of buffer A1 flowing through the collector outlet.
11. Take the first serum sample (labeled as XX-1-Y). Dilute the 20 µl serum sample with 80 µl PBS and spin at 13,000 rpm for 5 min to pellet insoluble debris.
12. End any run program.
13. Insert 1.5 ml microfuge tubes to position A1 to B10 (25 tubes) on the 12mm collector plate.
14. Draw up the 100 µl supernatant with 100 µl syringe and inject the sample into the loop. Do not remove the syringe from the injection port until the current run is finished.
15. Using the *Rohan\DepletionDualRunMethod* program (Method 1, Appendix 1) under depletion method saved on the AKTA purifier, start the run.
16. Pool the flow thru fractions A6 to B6 (8ml) into a 15 ml round-bottomed polyethylene tube (with cap). Place the appropriate label in the tube. Put on ice.
17. After the run is complete, remove both columns from the system
18. Disconnect the two columns
19. Connect the SuperMix (LC1) to the system. Add 1.5 ml tubes to the fraction collector at A1 to A15 (15 tubes).
20. Elute SuperMix (LC 1) with Buffer B1 by Method 2. (Appendix1)
21. Pool the fractions A4 to A9 (with OD_{280nm}), label the pool as EluteXX
22. Disconnect the SuperMix column from the system
23. Connect the Ig14 column (LC5) to the system
24. Start run Method 3 (Appendix1) to regenerate the column. Do not collect the eluent.
25. repeat steps 8 through 24 for serum sample labeled XX-2-Y
26. repeat steps 8 through 24 for serum sample labeled XX-PL-Y
27. Save all files for this triplet (case, control and pool) in the Rohan directory. The naming convention for these files should be:
    a. 'DP'<Stratum#><ID#><Aliquot#>
       Example: DP010101
       DP010201
       DP01PL01
       Eluent010101
       Eluent010201
       Eluent01PL01
       LC5-010101
       LC5-010201

28. When the process is finished, back up all files onto the Rohan portable hard drive in the DEPLETION directory.
29. Integrate the depletion curve to calculate the peak areas. Use this as quality control within and among stratum. Print out the integrated curves and the integrated results.

III. Trypsin Digestion (concentration, buffer exchange and trypsin digestion; process & digest the 6 samples from section 1 above simultaneously)

Buffers & reagents

5% Triton X-100 containing 0.02% sodium azide, store at room temperature.
50 mM TEAB (triethylammonium bicarbonate), freshly prepared from 1M stock.
TPCK treated trypsin, 20 µg /vial (Promega, Cat. # : V5111), store at -20°C.

Procedure

30. Pre-treat 12 Ultra-4 centrifugal concentrators with 4 ml 5% Triton X-100 for 30 min at room temperature (from step 3 in previous section), followed by extensive washing with distilled water to remove the Triton X-100
31. Transfer 4ml of each of the 8ml pooled fractions from the depletion steps to each designated concentrator prepared in step 30 (i.e. each sample will be concentrated in 2 concentrators).
32. Spin at 3,960 x g for 50 min at 8°C using a swinging bucket centrifuge to reduce the volume to 80 µl or less. This can be judged against the calibration line on the tube.
33. Dilute each concentrate with 2ml of 50 mM TEAB (triethylammonium bicarbonate buffer, Fluka, lot and filling code: 1192744 34005180). Pool the concentrates of each sample to one of its own concentrator (it will end up with 6 concentrate from 12), spin as in step 32.
34. Add 4 ml of 50 mM TEAB. Repeat spin as step 32, to reduce volume of sample to ≤40 µl.
35. Transfer the concentrated samples to a set of 500ul tubes (label the tubes first) separately. Wash the filter with 10 µl of TEAB buffer and transfer the washing to the tube. Repeat the wash once more with 5ul of TEAB buffer, collect to the tube.
36. Set the 100 µl pippet at 55µl. Draw up the sample in to the tip up to the maximum of the sample volume, compensate the difference with 50 mM TEAB to 55µl. Add 15µl acetonitrile to each tube thus to make the final sample containing 20% acetonitrile.
37. Dissolve each of 2 vials of trypsin in 20 µl of 50 mM TEAB. Each vial contains 20 ug TPCK treated trypsin (Promega, Cat: V5111).
38. Add 5µl of trypsin preparation to the sample each of the three flow thru samples and the SuperMix eluent samples from step 36. The total digestion volume for each sample is 75 µl.
39. Incubate at 37°C overnight at 400 rpm on a thermomixer.

IV. iTRAQ labeling (Label the 6 samples from section 2 above simultaneously, each with a unique iTRAQ reagents)

Buffers & reagents:

100% ethanol as supplied in the iTRAQ kit

iTRAQ reagents 2 set of 114, 115, 116 from the "iTRAQ Reagent Multi-Plex Kit" from Applied Biosystem. Store at -80°C.

Procedure

40. Spin down the digests resulted from step 39
41. Move two sets of iTRAQ reagents and ethanol to room temperature, wait for 20 min. Spin the reagents briefly. Make sure that the reagents are within storage period which is 12 months at -20C
42. Add 225ul 100% ethanol (absolute, HPLC grade, from iTRAQ kit) to each vial of iTRAQ reagent. Votex to dissolve the reagent completely. Spin briefly to collect all solution to the bottom.
43. Test the quality of the iTRAQ reagents by adding 3 ul of each labeling reagent to 1 µl of the 1:10 (in 50 mM TEAB) diluted 2 mix peptides (bradykinin-904 Da and ACTH-2465 Da), for 45 min at room temperature.
44. After test labeling, add 6 ul of H2O to each tube, mix. Take 3 µl of each into one 500 µl tube, add 9 µl of the matrix, Mix.
45. Detect the MS and MS/MS profile of the bradykinin using the 4800 instrument. If (1) in MS mode no 904 is observed while 1048 appears instead, and (2) in MS/MS mode of 1048 as the precursor ion, the product ions 114, 115, and 116 are labeled equally, continue the labeling of serum sample with the rest of the reagents as follows:.
46. Transfer 75µl of each digest to each corresponding iTRAQ reagent vial (114 reagent for ID 1 of each stratum, 115 reagent for ID 2 of each stratum, 116 for the pool sample for each stratum), vortex each tube to mix. Spin briefly. Incubate the labeling at room temperature overnight.
47. Stop the labeling by adding 900µl of SCXA1 buffer and allow excess reagent to completely hydrolyze over an additional 30 min at room temperature. Spin down briefly.
48. Save 20% (240ul) of labeled digest at -80C for glycoprotein determination.
49. Tansfer 960µl (out of 1200ul) of each to a 50 ml tube. Add 3.12ml SCX (strong cation exchange) buffer A (final total volume = 6ml).
50. Place the 'iTRAQ' label for the current stratum on the tube.
51. The combined and diluted sample (6ml) should be at pH 3.0 and can be stored at 4°C overnight. (It will contain 1 mM TEAB, 5.2mM potassium phosphate, 9% ethanol, 13% acetonitrile)

V. SCX-strong cation exchange chromatography (for the sample from the flow-through of both columns only)

Column

PolySulfoethyl A column (2.1x 100 mm, 5 µm, 300Å, volume≈0.3ml) from Poly LC Inc. Cat. No. 102 se0503

Guard column for the above

Buffers:

Buffer A: 25% acetontrile, 10 mM potassium phosphate, pH 3 (Made by titrating phosphoric acid with 5M KOH to pH 3.0)

Buffer B: 25% acetontrile, 10 mM potassium phosphate, pH 3, 700 mM KCl.

Instrumentation

AKTA purifier
Injection loop (8 ml)
Flow cell (Cat. 18-1147-25) (3 mm path length, 0.7 µl hold up volume)
(check the synchronization of UV tracing and sample collection of the fraction collector).
Syringe: syringe from Hamilton (10 ml), cat. No.: 7650-01; or 10 ml PP/PE lubricants free syringe (10 ml).

Run Method

See Appendix 2.

Procedure

52. Check that all solutions (A1 and B1) are ready. Check if the pH is correct, otherwise adjust it to pH 3.0. Filter if needed.
53. Turn on AKTA work-station and computer. Wait for system synchronization to complete.
54. Connect the 8 ml loop to the right position.
55. Connect UV light, pH meter, and Conductivity meter to the right position.
56. Run the *SCX equilibrate-system* program under Method, to bring all solutions to running-condition. (Important ! Manually purge **both** pumps in each delivery channel before running the equilibration program)
57. Set the flow rate to 0.5 ml/min. Test the flow rate with 1 ml per fraction for 3 fractions. Use a pipette to make sure the volume is 1ml. Test with pump A and pump B separately.
58. Set the pump rate to 1ml/min and allow 4ml of buffer B to flow through and then wash the flow path with 5ml of buffer A at 1ml/min. Finally set the pump rate to 0.1ml/min.
59. Switch valve to “inject” position to wash and purge the loop with 20 ml of Buffer A. Then switch valve to “load” position to wash and purge the loading syringe adaptor and the loop with 20 ml of Buffer A from a syringe fitted with the loading needle.
60. Connect the PolySulfoethyl A column (connect the guard column first) to the right position of the flow path.
61. Wash the column with 2 ml of buffer A at 0.1 ml/min.
62. Test the fraction collector tubing to make sure it’s not clogged.
63. Run a blank with the method to see if the gradient, monitored by conductivity curve, is the same as previous blank runs.
64. End any run program.
65. Turn off the power and start again. This will re-synchronize the pumps for a reproducible volume delivery at low speed.
66. Insert 1.5 ml tubes to collector rows A1 to A3, D1 to D16, E1 to E2, G1 to G16 and H1 to H10.
67. Draw up the sample from step #51 with syringe and load the sample into the loop. Do not remove the syringe from the injection port until the current run is finished.
68. Using the *Rohan\SCXRunMethod* method saved on the AKTA purifier, start the run. Collect the eluted peptide. Combine fractions D9 to E2 as ***F1***, G1 to G3 as ***F2***, G4 to G7 as ***F3***, G8 to G10 as ***F4***, G11 to H1 as ***F5***. There might be slight variation in combination method, depending on the SCX spectrum profile. Compare with previous SCX runs to decided the exact combination method. Store at 4C if not using immediately.

69. After the run, regenerate the column by running the column with a blank gradient followed by 6 ml of buffer A. Finally, wash the column and the system with at least 20 ml of 20% ethanol and store the column in 20% ethanol at room temperature.
70. Save the computer file generated by this process in the Rohan directory and on the portable hard drive in the iTRAQ subdirectory. The naming convention for these files should be:
    b. 'iTRAQ'<Stratum#><Aliquot#>
       Example: iTRAQ0101
71. Integrate the curve to calculate the elute peak areas. Set the X-axis window as 10 to 14 ml, Y-axis window as 0 to 1200 mAu at UV214. Re-set the baseline by drawing points. Print out the integrate results.
72. For samples from the SuperMix eluent, no SCX is needed.

VI. RPLC-reversed phase chromatography interfaced with Probot automatic spotting (for the SCX fractions and the eluate from SuperMix column)

Columns:

Trap column
    Dionex P/N 6720.0012
Loop
    250 µl
RP C18 column
    Dionex, P/N 160295, 300 um I.D.x, 15cm, static phase: C18, PepMap100, 3um, 100A
ABI 4800 Analyzer compatible 384 well MALDI insert plate

Solutions:

Solvent A
    5% acetronitrile, 0.1% TFA
Solvent B
    95% acetontrile, 0.1% TFA

Run Method

See Appendix 5 and 6

Procedure

73. Speed-Vac each SCX fraction to 98 ul. Adjust to 0.1% TFA by adding 1 ul of 10% TFA. Add 1 ul of LC internal calibration peptides.
74. Label 6 U3000 LC sampling vials as F1. F2, F3, F4, F5 and Eluent.
75. Transfer the Fractions to each designated LC vial.
76. Prepare 6 new 384 well TOFTOF plate, write on the plate the sample identity (RP4 F1, F2, F3, F4 F5 and Eluent). Plate preparation includes washing with 10% RBS/water solution, followed by washing with a substantial amount of water (many times and within each well to remove the detergent) and polishing (must make sure that all of the polish has been removed).
77. Spot the standard calibration MS/MS peptides in the 13 calibration spots of each plates.

78. Put two plates on the Probot platform each time.
79. Set the Dionex ultimate 3000 and the Probot at working condition (which includes analyzing the 16 peptide standard mixture and a blank to ensure that the U3000 is performing optimally).
80. Put the LC sample vials in Autosampler positions RA1 to RA6
81. For SCX F1 to F5, use the LC run method 5 in Appendix 1.
82. For Supermix column eluate, use the LC run method 6 in Appendix 1. In Eluent LC run peptides are spotted onto one 384 well plate (Eluent).
83. When each step is complete, save the computer files generated by this process in the Rohan directory and a second copy on the portable hard drive in the SCX directory. The naming convention for these files should be:
    c. 'SCX'_<Stratum#>_<Aliquot#>_'FR'<Fraction#>
    Example: SCX_01_1_F1
    SCX_01_1_F2
    .
    .
    SCX_01_1_F5
84. After all runs, regenerate the column by running the column with a blank gradient followed by 1 ml of buffer A. Finally, wash the column and the system thoroughly with 70% isopropanol and store the column in 70% isopropanol at room temperature.

VII. MS/MS (ABI 4800 MALDI-TOF/TOF for MS and MS/MS of iTRAQ labeled peptides)

85. Open the AB 4800 software Series Explorer software if it is not already running
86. Create a new spot set for each plate (BCPP_Stratumxx_SCX_Fx)
87. Insert one plate in to the autoloader after blowing nitrogen on the plate
88. Open the MS and MS/MS Acq Methods and their corresponding Proc Methods.
89. Load the created spot set after selecting the Load Plate icon  and select the appropriate plate name
90. Select the MS Acq & Processing Method and set each one as Active by clicking on the Set as Active icon  .
91. Select any of the 13 Calibration spots and start acquisition by clicking  on the toolbar or selecting Interactive > Start Active Acquisition Method**.** Note – If the laser was not on it may take at least 2 minutes for it to turn on. A live trace will be displayed in the Spectrum Viewer along with its final trace**.**
92. Acquire from 3 well (1 each from row B, E & O) MS1 data and note at least one peptide for subsequent MS/MS.
93. Align the loaded sample plate by selecting Plate > Align Sample Plate and follow the directions from the Sample Plate Alignment Wizard. Improper alignment of the plate may cause some sample spots to be missed by the laser.
94. Calibrate the sample plate with the MS reflector acquisition method as follows:

a. Move to one of the 13 calibration spots and start acquisition
b. Check the quality of the spectra and correct if necessary
c. Select the MS reflector processing method, click on the Calibration tab, select Internal as the Calibration Type and click on save
d. Select the Spot Set Tab in the Spot Set Window and in Spot Set Manager select spots CAL1-CAL13 (click and drag across the spots)
e. Right mouse click on the highlighted spots and select Copy Spot Sets to Job using Run Specific Methods
f. Select Plate Model and Default Calibration for the calibration type
g. Select MS reflector as the Acq Method & MS reflector as the Proc Method (the Cal Type will state internal for all 13 Cal spots)
h. Submit the Spot Set Job to the queue by clicking on  in the toolbar or Select Batch > Submit Spot Set to Job. The software will automatically validate the submitted spot set job
i. Start the Job Queue by clicking on  in the toolbar or Select Batch > Start Job Queue. Examine the spectrum of the calibration standard and check that the peaks of interest are present, masses are within the expected range, the peaks are narrow and well resolved and that the signal is not saturated (> 9.0E+4). If the signal or resolution is below optimum contact either Eddie or Hui.

95. Calibrate the sample plate with the MS/MS reflector acquisition method as follows:
a. Move to one of the 13 calibration spots and start acquisition
b. Check the quality of the spectra and correct if necessary
c. Select the MS/MS reflector processing method, click on the Calibration tab, select Internal as the Calibration Type and click on save
d. Select the Spot Set Tab in the Spot Set Window and in Spot Set Manager select spots CAL1-CAL13 (click and drag across the spots)
e. Right mouse click on the highlighted spots and select Copy Spot Sets to Job using Run Specific Methods
f. Select None for the calibration type
g. Select MS/MS reflector as the Acq Method & MS/MS reflector as the Proc Method (the Cal Type will state internal for all 13 Cal spots)
h. Submit the Spot Set Job to the queue by clicking on  in the toolbar or Select Batch > Submit Spot Set to Job. The software will automatically validate the submitted spot set job
i. Start the Job Queue by clicking on  in the toolbar or Select Batch > Start Job Queue. Examine the spectrum of the calibration standard and check that the peaks of interest are present, masses are within the expected range, the peaks are narrow and well resolved and that the signal is above . If the signal or resolution is below optimum contact either Eddie or Hui.

96. Select Default as the calibration type for both MS & MS/MS reflector processing files and save the files.

97. Select the well position chosen in step #94 and type in the peptide mass in the precursor option in the MS/MS refector Acq Method. Start acquisition and check that the laser power is optimized for the iTRAQ reporter and peptide sequence ions. Gradually decrease or increase the laser power for optimization.

98. Select all the sample wells of the sample plate to be analyzed in the Spot Set Tab (A1-P1) and do not include the calibration spots.

99. Right mouse click on the highlighted spots and select Copy Spot Set to Job > Using Run Specific Methods
100. For the Calibration Type select None
101. Select the MS Reflector Acquisition and Processing Methods (ensure that all the rows have been filled down).
102. Choose JobWide Interpretation and select your JobWide Interpretation Method.
103. Check that the acquisition, processing and jobwide interpretation methods are loaded for every spot.
104. Submit the Spot Set Job to the queue by clicking on  in the toolbar or Select Batch > Submit Spot Set to Job. The software will automatically validate the submitted spot set job
105. Start the Job Queue by clicking on  in the toolbar or Select Batch > Start Job Queue.
106. <u>After each plate's analysis, draw the 116 signal dot-plot to see the distribution of 116 reporter against retention time.</u>

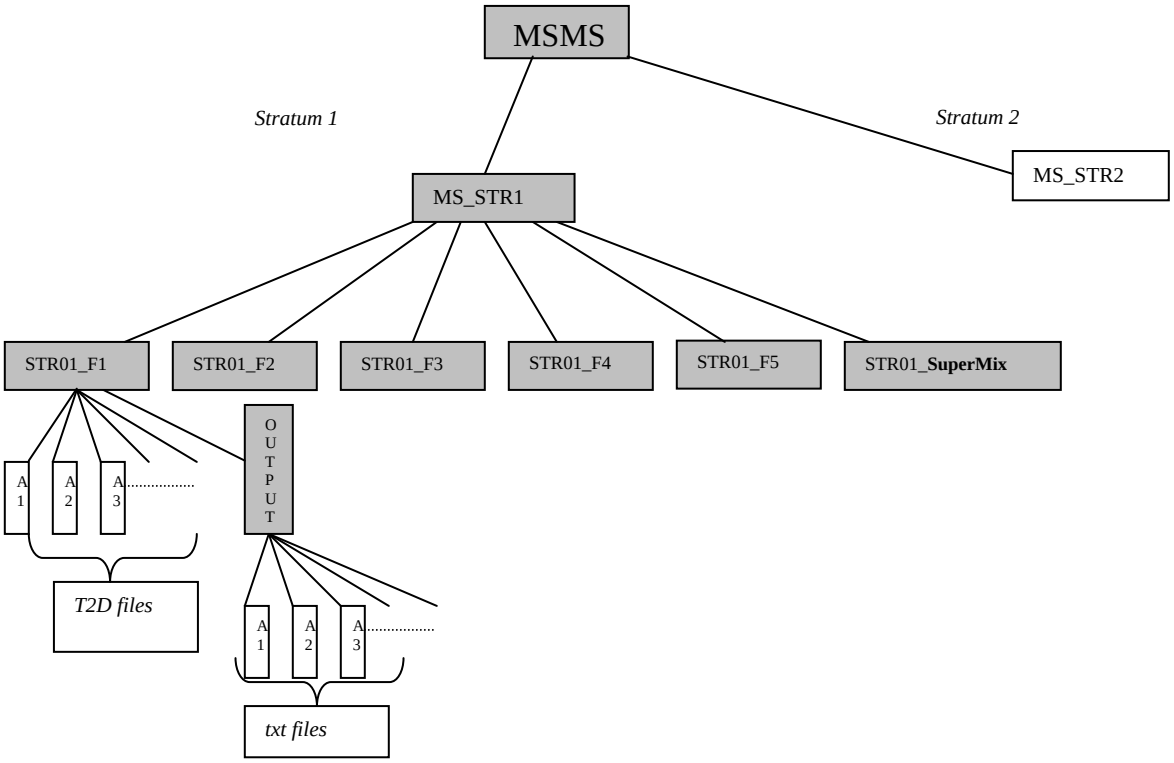
VIII. Data Management and Analysis (Protein ID and label ions comparison)

In this step, the spectrum will be analyzed and the intensities of the iTRAQ reporter ions will be compared. At the conclusion of the MSMS run for each stratum, the following procedures will be run.

Procedure

1. Open the spotset, display the MS/MS run, highlight all rows from A1 to P1, right mouse click on the highlighted rows, select export to folder, select or create the corresponding folder name and export as T2D files. Repeat for all the SCX acquired plates and copy to the portable hard drive into the appropriate directory for that stratum.
2. After a complete run on a stratum is complete and all T2D files have been copied to the portable hard drive (approximately once every week to 10 days) , the DEPH will copy the hard drive data onto their servers, delete all data on the hard drive and return the hard drive to the lab so that analysis on the next stratum can begin.

3. A figure representing the directory structure for each stratum appears below*: (Directories are shaded)*



4. A program will be run on all .t2d in a given fraction directory to convert them to text files for analysis. Analysis files will be created for each fraction and they will be merged into one analysis file for each stratum (see below) which will contain precursor information. The program Protein Pilot will be used to obtain protein ID and quantitation.

not normalized

Fraction #	Spot #	Elution Time	M/Z (peptide mass data)	Intensity for 114	Intensity for 115	Intensity for 116	Base	SD of Base
F1	A1							
F1	.							
F1	.							
F1	B1							
F1	.							
F1	.							
F1	.							
F1	B1							
F2	A1							
F2	.							
F2	.							
F2	B1							

F2								
F2	.							
F3	.							
.	.							

5. When analysis of all stratum data are complete, a program created by the proteomics lab will be run on these stratum files to create the final analysis file which will contain the following data:

Peptide Mass	Retention Time (Spot #)	Stratum 01 (114) Intensity After Normalization	Stratum 01 (115) Intensity After Normalization	Stratum 01 (116) Intensity After Normalization	Stratum 02 (114) Intensity After Normalization	Stratum 02 (115) Intensity After Normalization	Stratum 02 114 Intensity After Normalization		Stratum 40 116 Intensity After Normalization

IX. Appendix:

1. Method1, Depletion Run
2. Mothod2, SuperMix (LC1) elution collection regeneration
3. Method3, Ig14 (LC5) regeneration
4. Method of SCX Run
5. Method of LC Run for SCX fractions
6. Method of LC Run for SuperMix eluate

1. Method1, Depletion Run

   Buffer A1: Phosphate Buffered Saline (PBS), pH 7.4
   Buffer B1: 0.1 M glycine-HCl, pH 2.5
   Buffer A2: 0.1M Na_2HPO_4-NaH_2PO_4, pH 8.0
   Flow rate: 0.5 ml/min
   0.00 Valve at Inject position to load the sample from loop to columns
   0.00 100% A1 for 26 min
   0.00 Start collection. 0.5 ml/fraction. Collect A1 to B15.
   26.00 Valve at Load
   26.00 Stop run.

2. Method2, SuperMix Elution

   Buffer A1: Phosphate Buffered Saline (PBS), pH 7.4
   Buffer B1: 0.1 M glycine-HCl, pH 2.5

Buffer A2: 0.1M Na_2HPO_4-NaH_2PO_4, pH 8.0
Flow rate: 2 ml/min

0.00 Load position
0.50 100% Buffer B1 for 2.5 min
0.50 Start collection at 0.5 ml/fraction, from A1 to A15.
3.00 100% A2 for 2 min
5.00 100% A1 for 3 min
8.00 End run.

3. Method3, Ig14 (LC5) regeneration
   Buffer A1: Phosphate Buffered Saline (PBS), pH 7.4
   Buffer B1: 0.1 M glycine-HCl, pH 2.5
   Buffer A2: 0.1M Na_2HPO_4-NaH_2PO_4, pH 8.0
   Flow rate: 2 ml/min

   0.00 Load position
   0.00 100% Buffer B1 for 13 min
   13.00 100% A2 for 5 min
   18.00 100% A1 for 24 min
   42.00 End run.

4. Method of Strong Cation Exchange (SCX) Run
   Solution A1: 10 mM potassium phosphate, pH 3.0, 20% acetonitrile
   Solution B1: 10 mM potassium phosphate, pH 3.0, 700 mM potassium chloride, 20% acetonitrile
   0.00 Valve at Load position, inject the sample to loop
   0.00 Equilibration the column, 100% Solution A1, for 15 min at 0.1ml/min
   15.0 Valve at Inject position (loading column), 100% Solution A1, for 100 min at 0.1 ml/min
   115.0 Wash with A1 for 6 min at 0.1 ml/min
   121.0 Start gradient B1 from 0% to 10% in 2 min at 0.1ml/min
   123.0 Gradient B1 from 10% to 25% in 17 min at 0.1ml/min
   140.0 Gradient B1 from 25% to 100% in 5 min at 0.1ml/min
   145.0 Gradient B1 at 100% for 26 min at 0.1ml/min
   171.0 Gradient B1 at 0%
   240.0 End of Run

5. Method of LC-Spotting Run for SCX fractions
   Solution A: 5% acetonitrile. 0.1% TFA, 95% H2O
   Solution B: 95% acetonitrile. 0.1% TFA, 5% H2O
   Flow rate: MicroPump at 4µl/min constant. LoadingPump at 40µl/min constant.
   0.00 Load trap column with LoadingPump at 40µl/min for 30 min, ValveLeft at 10-1 position, 5%B, 95%A
   30.0 ValveLeft at 1-2 position, MicroPump at 4µl/min, for 10 min Gradient of B from 5% to 15%
   35.0 Probot turns on. Collect 50 min.

40.0 %B from 15 to 25.5 in 35 min,
75.0 %B from 25.5 to 90 in 10 min,
85.0 %B hold at 90 for 5 min,
90.0 %B from 90 to 5 for 10 min,
Probot off,
100.0 %B hold at 5 for 10 min,
110.0 LoadingPump flow at 0µl/min, MicroPump at 0 µl/min, ValveLeft turns to 10-1 POSOTION
110.1 End of Run

6. Method of LC-Spotting Run for SuperMix eluate

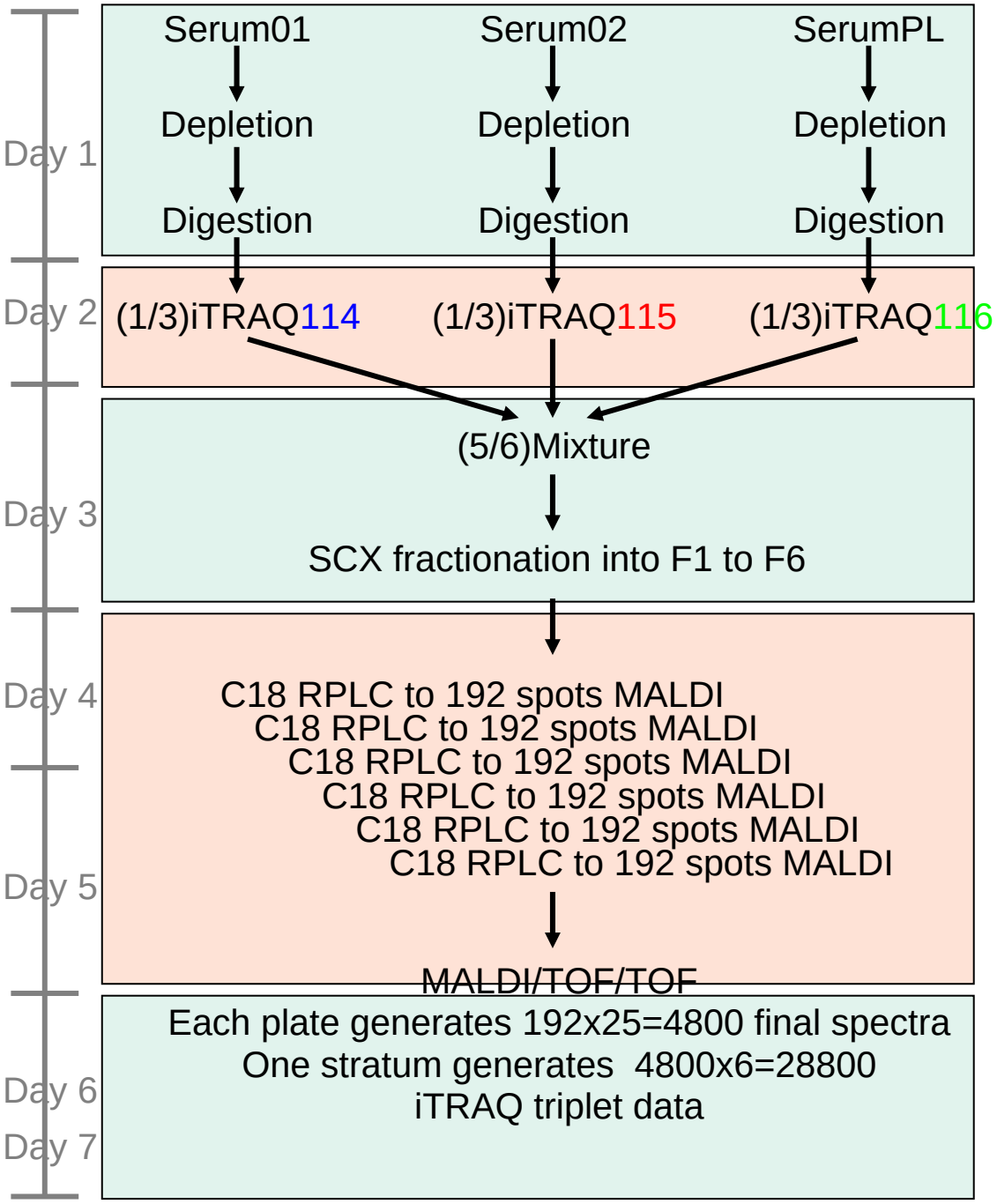
Solution A: 5% acetonitrile. 0.1% TFA, 95% H2O
Solution B: 95% acetonitrile. 0.1% TFA, 5% H2O
Flow rate: MicroPump at 4µl/min constant. LoadingPump at 40µl/min constant.
0.00 Load trap column with LoadingPump at 40µl/min for 30 min, ValveLeft at 10-1 position, 5%B, 95%A
30.0 ValveLeft at 1-2 position, MicroPump at 4µl/min, for 10 min Gradient of B from 5% to 15%
35.0 Probot turns on. Collect 50 min.
40.0 %B from 15 to 30 in 50 min,
90.0 %B from 30 to 90 in 5 min,
95.0 %B hold at 90 for 15 min,
110.0 %B from 90 to 5 for 5 min,
Probot off,
115.0 %B hold at 5 for 10 min,
125.0 LoadingPump flow at 0µl/min, MicroPump at 0 µl/min, ValveLeft turns to 10-1 POSOTION
125.1 End of Run

Appendix 2

Grand Check List

StratumXX
022608

Work Flow Chart



Depletion/Digestion Run Check List

StratumXX

022608

- [ ] 1. Prepare the following solutions freshly
  - Buffer A1: PBS. Dissolve 1 bag of sigma PBS powder in H2O and make up to 1000 ml with H2O
  - Buffer B1: Stripping buffer: Add 20 ml of 10X stripping buffer to 180 ml of H2O, mix
  - Buffer A2: Neutralization buffer: Add 20 ml of 10X neutralization buffer to 180 ml of H2O, mix
- [ ] 2. Wash the pump heads with 20% ethanol
- [ ] 3. Turn on AKTA system and computer
- [ ] 4. Connect UV cell, conductivity meter to the correct position
- [ ] 5. Wash the pH meter tip with H2O, connect to the pH cell
- [ ] 6. Connect the correct loop (250 ul) to the correct position, wash the injection hole and check the hole size for compatibility with syringe needle
- [ ] 7. Connect the collection tubing directly to the collection tip, bypassing the accumulator.
- [ ] 8. Make sure that the collection path is not clogged. Wipe the Syncdrop with H2O to clean up any stain on its walls (if the walls are dirty, the SyncDrop message show up and collection will not start). Make sure that the tip is in the middle and is tightly screwed in.
- [ ] 9. Equilibrate the tubing and pumps with 10 ml of H_2O for each of the 4 pumps at 2 ml/min
- [ ] 10. Wash pumps A and B with H_2O, using the built in protocol (PumpWashBasic)
- [ ] 11. Wash the loop with 4 ml of H_2O at 4 ml/min
- [ ] 12. Stop flow if it is on
- [ ] 13. Replace Inlet A1 liquid with Buffer A1, B1 liquid with BufferB1, and A2 liquid with Buffer A2, degas for 5 min
- [ ] 14. Start a DepletionEquiSystem method, to bring solutions up to Outlets
- [ ] 15. Wash the loop with 4 ml of Buffer A1 at 1 ml/min, switch from Load to Injection position twice to remove any trapped air bubble
- [ ] 16. Set flow rate to 0.2 ml/min
- [ ] 17. Immediately after moving the depletion column from 4ºC to room temperature, connect the column to the correct position following the correct flow direction.
- [ ] 18. Swiftly purge the column with Buffer A1 at 1 ml/min, watch the pressure (normal pressure=0.1)
- [ ] 19. Connect the column outlet to the system, watch the pressure (normal pressure=0.1)
- [ ] 20. Run 10 ml of Buffer A1 through the column, watch the pressure (0), UV(50 at 214), pH(7.4), Conductivity(?)
- [ ] 21. Add 1.5 ml tubes to the 12mm collector rack at A1 through A9
- [ ] 22. Thaw the serum samples on ice
- [ ] 23. Wash the 100 ul syringe with Buffer A1 3 times, 100 ul each, wipe clean the needle with Kimwipe
- [ ] 24. Draw 100 ul of Buffer A1 to the syringe, purge the air bubbles
- [ ] 25. Assure that the valve is at Load position
- [ ] 26. Inject the syringe content to the loop slowly. Make sure there is no leak at the needle tip and observe the liquid drop in the waste tank. Leave the syringe there
- [ ] 27. Double-check the above items
- [ ] 28. Make a blank depletion run with Buffer A1 that is injected
- [ ] 29. Go to File\Run\Rohan\DepletionRunMethod, click Start, observe the whole run especially at every check point.
- [ ] 30. Most importantly, the collector is on and the collection starts at the right time and in the right tubes. Observe the spectrum
- [ ] 31. After run, examine that each fraction contains 500 ul of run thru
- [ ] 32. Discard the collections from the blank run. Change to new tubes
- [ ] 33. Spin down the Stratum38-PL-2
- [ ] 34. Add 80 ul of Buffer A1, mix and spin at full speed (13,000 rpm) in a microfuge for 5 min
- [ ]

- [ ] 36. with the 100 ul syringe, take 100 µl. Wipe the needle
- [ ] 37. inject the 100 ul of serum to the loop
- [ ] 38. leave the syringe at inject position
- [ ] 39. Go to File\Run\Rohan\DepletionMethodRun, click Start, observe the whole run especially at every check point.
- [ ] 40. Most importantly, the collector is on and the collection starts at the right time and in the right tubes. Observe the spectrum
- [ ] 41. After run, examine that each fraction contains 500 ul of run thru
- [ ] 42. collect the fractions from A3 to A8, including A3 and A8 (total is 2.5ml)
- [ ] 43. designate the tube as DP 090101
- [ ] 44. Label three spin ultrafiltration concentrators as conc05-1-2, Conc05-2-2 and Conc05-pool-2, separately.
- [ ] 45. Add 4 ml of 5% Triton-X100 (5k cut off) to each tube. Leave them on bench for more than 30 min
- [ ] 46. Spin down Stratum 39-PL-2
- [ ] 47. Add 80 ul of Buffer A1, mix and spin at full speed for 5 min
- [ ] 48. with the 100 ul syringe, take 100 ul. Wipe the needle
- [ ] 49. inject the 100 ul of serum to the loop
- [ ] 50. leave the syringe the at inject position
- [ ] 51. Go to File\Run\Rohan\DepletionMethodRun, click Start, observe the whole run especially at every check point.
- [ ] 52. Most importantly, the collector is on and the collection starts at the right time and in the right tubes. Observe the spectrum
- [ ] 53. After run, examine that each fraction contains 500 ul of run thru
- [ ] 54. Collect the fractions from A3 to A8, including A3 and A8 (total is 2.5ml)
- [ ] 55. Designate the tube as Dep 09-2-1
- [ ] 56. Repeat steps 34 to 42 for sample 09PL01
- [ ] 57. Designate the tube as DP09PL01
- [ ] 58. Spin down the three collection tubes
- [ ] 59. Rinse the three spin concentrator from step (45) thoroughly with H_2O
- [ ] 60. Transfer the 3 Dep samples to the corresponding concentrators
- [ ] 61. Spin the concentrators (+samples) at 4000g for 40 min first (vol. reduce to 80 ul); mix in 3 ml of 50 mM TEAB buffer.
- [ ] 62. Spin again at 4,000g for 40 min. Volume must reduce to less than 80ul. Spin once more if needed.
- [ ] 63. Transfer the concentrates individually to three 500-ul tubes labeled iTRAQ090101, 090201, and 09PL01
- [ ] 64. If the volume in the any tube is less than 80, add 50 mM TEAB buffer to make up the volume to 80 ul
- [ ] 65. Add 20 ul of 100% ACN to each tube
- [ ] 66. Dissolve one vial of trypsin in 20 ul of 50mM TEAB with Add 5.5 ul of this reconstituted trypsin solution to each sample
- [ ] 67. Vortex and spin down
- [ ] 68. Incubate at 37ºC overnight, shaking at 400rpm.
- [ ] 69. Store the digest at 4ºC if not labeling directly.
- [ ]

Day 2 continues on next page

Depletion Run Check List (page 3), Sample: StratumXX

- ☐ 71. From -80C freezer Box A3.3, Move iTRAQ reagent (Lot No. 0710076, Foster City, CA 94404) 114, 115 and 116, and 100% EtOH to room tesmprature, leave on bench for 10 min, spin briefly (10 sec in a microfuge) to collect the content to the bottom of the tubes.
- ☐ 72. Add swiftly 105 ul of EtOH to each vial, vortex for 30s. Spin down briefly.
- ☐ 73. Dilute 6 ul of the stock dual standard peptide (des-Arg1-Bradykinin Mr=904.5, ACTH18-39, Mr= 2465.2) with 24 ul of 50 mM TEAB, mix well.
- ☐ 74. Prepare three 500ul tubes, label them as pep114, pep115 and pep116. Add 3 ul of iTRAQ reagents to each corresponding tube
- ☐ 75. Add 1 ul of the diluted standard peptide to each tube, mix and spin
- ☐ 76. Incubate at room temperature for 45 min
- ☐ 77. Spin the digests briefly (10 sec. in a microfuge), add 6 ul of H2O to stop reaction.
- ☐ 78. Take 3 ul from each into a new 500 ul tube, mix, add 9 ul of ready matrix, mix
- ☐ 79. Load to spot on 4700 plate, do MS and MSMS scan to determine the labeling efficiency and the distribution of the reporter ions.
- ☐ 80. If labeling test passes, contunue the labeling with the serum digest by add 40 ul of digest to each designated iTRAQ vial. Mix, spin, incubate overnight.
- ☐ 81. Store the EtOH at -80ºC with the Kit in Box A1.1.
- ☐ 82. Stop the reaction by add 420 ul of H2O to each tube (total 560 ul in each tube), spin to pellet the precipitation
- ☐ 83. Take 466 ul of each to a 50 ml tube
- ☐ 84. Add 4,300 ul of SCXA1 solution , mix gently, test pH with paper. pH should be at 3.0
- ☐ 85. Save the rest of 94 ul at -80ºC freezer Box C2.3. at position A5, A6, A7.
- ☐ 86. Subject the diluted mixture to SCX fractionation as the following

SCX Run Check List

Sample__StratumXX
Date_____020608

- ☐ 1. Purge the pump heads with 25% ethanol using a 50 ml syringe to remove air bubble
- ☐ 2. Degas the H_2O for more than 5 min
- ☐ 3. Switch inlets to H_2O
- ☐ 4. Connect UV cell, conductivity meter to the correct position
- ☐ 5. Wash the pH meter tip with H2O, connect to the pH cell
- ☐ 6. Connect the correct loop to the correct position, wash the injection hole and check the hole size for compatibility with syringe needle
- ☐ 7. Connect the collection tubing directly to the collection tip, bypassing the accumulator.
- ☐ 8. Make sure that the collection path is not clogged. Wipe the Syncdrop with H2O to clean up any stain on its walls (if the walls are dirty, the SyncDrop mesagse show up and collection will not start). Make sure that the tip is in the middle and is tightly screwed in.
- ☐ 9. Equilibrate the tubing and pumps with 10 ml of H_2O for each of the 4 pumps at 2 ml/min
- ☐ 10. Wash pumps A and B with H_2O, using the built in protocol (PumpWashBasic)
- ☐ 11. Wash the loop with 4- folds loop size (8 ml) of H_2O at 4 ml/min
- ☐ 12. Stop flow if it is on
- ☐ 13. Replace Inlet A1 liquid with SCXA1 and B1 liquid with SCXB1, degas for 5 min
- ☐ 14. Start a SCXEquiSystem method, to bring solutions up to Outlets
- ☐ 15. Wash the loop with 40 ml of SCXA1 at 4 ml/min, switch from Load to Injection position twice to remove any trapped air bubble
- ☐ 16. Set flow rate to 0.1ml/min
- ☐ 17. Connect the SCX Precolumn, watch the pump pressure (0.9 normal pressure)
- ☐ 18. Connect the SCX polyA column to the precolumn, watch the pressure (3.2 to 3.5 normal pressure)
- ☐ 19. Connect the column outlet to the system, watch the pressure (3.2 to 3.5 normal pressure)
- ☐ 20. Run 2 ml of SCX A1 through the columns, watch the pressure, UV, pH, Conductivity
- ☐ 21. Add 1.5 ml tubes to the following position on the 12mm collector rack: A1, A2, A3, D1 to D15, E1, E2, G1 to G15, H1 to H9
- ☐ 22. Bring the iTRAQ labeled digest mixture to room temperature, vortex, spin 1 min at full speed
- ☐ 23. Collect the 450 ul supernatant to 4 ml of SCXA1 in a 50ml tube, mix well gently
- ☐ 24. Wash the 10 ml syringe with SCX A1 3 times, 6 ml each, wipe clean the needle with Kimwipe
- ☐ 25. Draw all 4.50ml sample to the syringe, purge the air bubbles
- ☐ 26. Assure that the valve is at Load position
- ☐ 27. Inject the sample to the loop, slowly but firmly. Make sure there is no leak at the needle tip and observe the liquid drop in the waste tank. Leave the syringe there
- ☐ 28. Double-check the above items
- ☐ 29. Go to File\Run\SCX, click Start, observe the whole run especially at every check point.
- ☐ 30. Most importantly, the collector is on and the collection starts at the right time and in the right tubes. Observe the conductivity change when elution begins
- ☐ 31. After run, examine that each fraction contains 100 ul of elute
- ☐ 32. Pool fractions D8 to D13 into the D8 tube. Label the vial as S05B_F1
- ☐ 33. Pool fractions D14 to E1 into the E2 tube. Label the vial as S05B_F2
- ☐ 34. Pool fractions E2 to G2 into the G3 tube. Label the vial as S05B_F3
- ☐ 35. Pool fractions G3 to G6 into the G6 tube. Label the vial as S05B_F4
- ☐ 36. Pool fractions G7 to G10 into the G9 tube. Label the vial as S05B_F5
- ☐ 37. Pool fractions G11 to H3 into the G12 tube. Label the vial as S05B_F6
- ☐ 38. Store at 4°C for next day's SpeedVac processing.

Pooling according to absorbance at 214

U3000 LC Run Check List

If necessary (or every 2 weeks), Clean up the Tof-Tof at this step.

Sample_StratumXX
Date__010208____
Time_____

- ☐ 1. Move SCX fractions from 4°C to room temperature
- ☐ 2. Spin down drops on tube side
- ☐ 3. Prepare the SpeedVac, wait until the vacuum reach the proper level (0.1 atm)
- ☐ 4. Place the tubes balanced in the centrifuge
- ☐ 5. Start Drying process. Stop and check level of sample every 10 min until the sample volume is less than 100 ul
- ☐ 6. Reconstitute the sample volume to 100 with H_2O using a 200 ul pipetman
- ☐ 7. Add to each fraction 1 ul of LC standard peptides (peptide 1900, middle, and peptide 1826, late, -80C, BOX C2.1.) which are reconstituted in 20 ul of 50% ACN and 0.1%TFA
- ☐ 8. Add to each fraction 1 ul of 10% TFA, mix and spin
- ☐ 9. Label 6 U3000 autosampler 250ul vials as SCX05BF1 to SCX05BF6
- ☐ 10. Transfer the fractions to corresponding vial, tape the vial gently to remove trapped air bubbles in vials
- ☐ 11. Place the vials to designated positions in autosampler rack
- ☐ 12. Take 1 vial of ready to use Agilent CHCA matrix (6 mg/ml). Add 3 ul of TFA, mix.
- ☐ 13. Transfer 2 ml to the matrix bottle of the uCarrier dosage unit
- ☐ 14. Add 8 ul of the MS internal standard peptides (bradykinin and ACTH), mix by pipeting
- ☐ 15. Place the matrix bottle in the bottle holder in the dosage unit.
- ☐ 16. Manually purge the syringe, the valves and the spotting needle with the matrix by 5 strokes, be sure no air bubble in the syringe and be sure to switch the valve position prior to each operation
- ☐ 17. Place six disposable plate inserts in the uCarrier plate holder, write down the plate number in correct order
- ☐ 18. Align the well positions
- ☐ 19. Open the application file, and start application
- ☐ 20. Notice the message at the bottom of the application file as “waiting for the signal from the closure”
- ☐ 21. Prepare fresh solutions A and B for HPLC
- ☐ 22. Purge the pumps
- ☐ 23. Change a new trap column
- ☐ 24. Fill the autosampler wash liquid with solution A, be sure no air bubble in the autosampler syringe
- ☐ 25. Fill the R1 bottle with solution A (used for blank runs)
- ☐ 26. Fill the valve washing liquid bottle with H_2O
- ☐ 27. Turn on UV 214 and 254 nm, turn on data acquisition, watch the UV reading (650 mAu at 214 nm)
- ☐ 28. Set flow rat at 40 ul/min for loading pump and 2.0 ul/min for micropump
- ☐ 29. Observe the pressures (90 Value)
- ☐ 30. Start a blank run with R1 solution, 95ul ulPickup injection mode, observe the background
- ☐ 31. Estimate the flow rate at the outlet of UV meter (650 mAu at 214 nm)
- ☐ 32. Connect the outlet of UV to the uCarrier tubing, make sure that liquid comes out of the spotting needle and the pump pressures are not changed
- ☐ 33. Open the sequence editing page, set up a blank run before each sample run
- ☐ 34. Name the blank runs as blankR1, and the sample runs as LC-SC05BF1 to LC_SC05BF6
- ☐ 35. Make sure the correct vial positions for blank and samples are selected
- ☐ 36. Make sure set injection volume to 95ul
- ☐ 37. Make sure the correct program files are chosen for blank and samples
- ☐ 38. Save the sequence in a new folder under Rohan folder. Name the new folder as LC-SCX05B. Name the sequence file as LC-SCX05B, too
- ☐ 39. Final check: no liquid leak, LC-Spotting tubing connection, nothing block the uCarrier platform from moving, uCarrier power on, syringe filled with matrix, plates are in position
- ☐ 40. Start batch, watch any warning message, OK to start
- ☐ 41. Check the run periodically

MS/MS analysis check list

Stratum__StratumXX
Date__010808__

1. Before inserting the plate into the autoloader, record the plate bar code and the corresponding fraction and slot numbers.
2. Create a spot set for each plate in LC/MALDI mode
3. Map the spot set to the corresponding plate number and slot number.
4. Load plate to source chamber to align the plate position
5. Test the acquisition and processing method. First test the calibration standards, make sure that both MS1 and MS2 calibration pass the criteria. Then test the sample, spots B10, D10, F10 and H10. Take 3 MS1 peaks (representing high, medium, and low intensity peaks) for MS2 scan. If total intensity is low, increase the laser power.
   - MS1: yansen\BCPP\Methods\MS1_Positive_02May2007
   - MS1: yansen\BCPP\Methods\MS1_Default_Processing_02May2007
   - MS2: yansen\BCPP\Methods\MS2_Positive_02May2007
   - MS2: yansen\BCPP\Methods\MS2_Default_Processing_02May2007
6. Choose the methods from yansen\BCPP\Methods for the automatic runs.
7. Submit the Spot Set Job to queue run, observe the validation process.
8. Switch the acquisition mode from interactive to queue run mode.
9. Start the queue run.
10. Observe the beginning of the automatic run, pay attention to the signals.

Data Processing Check List

StratumXX

1. After the TOF/TOF scan, select 116 signals to draw a dot-plot against retention time. If 116 distribution doesn't represent LC spectrum, repeat scanning the abnormal area.
2. Using the ABI data management tool, export the selected ms/ms data files to the drive D:yansen\StratumXX.. as T2D files. Export repeat runs as well.
3. Copy the exported files to the external drive provided by the Epidemiology Dept.
4. Backup the data files to another external hard drive for data storage

Appendix 3

Proteins identified (based on hit number of 500)

Rank Protein Name
1 (P01024) Complement C3 precursor [Contains: C3a anaphylatoxin]
2 (P01023) Alpha-2-macroglobulin precursor (Alpha-2-M)
3 (P01028) Complement C4 precursor [Contains: C4a anaphylatoxin; C4b]
4 (P00450) Ceruloplasmin precursor (EC 1.16.3.1) (Ferroxidase)
5 (P02790) Hemopexin precursor (Beta-1B-glycoprotein)
6 (P04114) Apolipoprotein B-100 precursor (Apo B-100) [Contains: Apolipoprotein B-48 (Apo B-48)]
7 (P02751) Fibronectin precursor (FN) (Cold-insoluble globulin) (CIG)
8 (P19823) Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy chain H2) (Inter-alpha-i
9 (P00751) Complement factor B precursor (EC 3.4.21.47) (C3/C5 convertase) (Properdin factor B) (Glyc
10 (P04217) Alpha-1B-glycoprotein precursor (Alpha-1-B glycoprotein)
11 (P06727) Apolipoprotein A-IV precursor (Apo-AIV) (ApoA-IV)
12 (Q14624) Inter-alpha-trypsin inhibitor heavy chain H4 precursor (ITI heavy chain H4) (Inter-alpha-i
13 (P01042) Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains: Bradykinin (Kallidin I
14 (P08603) Complement factor H precursor (H factor 1)
15 (P01011) Alpha-1-antichymotrypsin precursor (ACT)
16 (P02774) Vitamin D-binding protein precursor (DBP) (Group-specific component) (Gc-globulin) (VDB)
17 (P01008) Antithrombin-III precursor (ATIII) (PRO0309)
18 (P19827) Inter-alpha-trypsin inhibitor heavy chain H1 precursor (ITI heavy chain H1) (Inter-alpha-i
19 (P20742) Pregnancy zone protein precursor
20 (P04196) Histidine-rich glycoprotein precursor (Histidine-proline rich glycoprotein) (HPRG)
21 (P00747) Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin]
22 (P04004) Vitronectin precursor (Serum spreading factor) (S-protein) (V75) [Contains: Vitronectin V6
23 (P43652) Afamin precursor (Alpha-albumin) (Alpha-Alb)
24 (P01031) Complement C5 precursor [Contains: C5a anaphylatoxin]
25 (P02753) Plasma retinol-binding protein precursor (PRBP) (RBP) (PRO2222)
26 (P02743) Serum amyloid P-component precursor (SAP) (9.5S alpha-1-glycoprotein)
27 (P00736) Complement C1r subcomponent precursor (EC 3.4.21.41) (Complement component 1, r subcompone
28 (P00734) Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II)
29 (P05546) Heparin cofactor II precursor (HC-II) (Protease inhibitor leuserpin 2) (HLS2)
30 (P02775) Platelet basic protein precursor (PBP) (Small inducible cytokine B7) (CXCL7) (Leukocyte-de
31 (P01019) Angiotensinogen precursor [Contains: Angiotensin I (Ang I); Angiotensin II (Ang II); Angio
32 (P02768) Serum albumin precursor
33 (P02749) Beta-2-glycoprotein I precursor (Apolipoprotein H) (Apo-H) (B2GPI) (Beta(2)GPI) (Activated
34 (P03952) Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein) (Kininogenin) (Fletcher
35 (P02765) Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Alpha-2-Z-globulin) (Ba-alpha-2-glycoprotein
36 (P02748) Complement component C9 precursor
37 (O75636) Ficolin 3 precursor (Collagen/fibrinogen domain-containing protein 3) (Collagen/fibrinogen
38 (P02760) AMBP protein precursor [Contains: Alpha-1-microglobulin (Protein HC) (Complex-forming glyc
39 (P04003) C4b-binding protein alpha chain precursor (C4bp) (Proline-rich protein) (PRP)
40 (P06396) Gelsolin precursor (Actin-depolymerizing factor) (ADF) (Brevin) (AGEL)
41 (Q06033) Inter-alpha-trypsin inhibitor heavy chain H3 precursor (ITI heavy chain H3) (Inter-alpha-i
42 (P02746) Complement C1q subcomponent, B chain precursor
43 (P00738) Haptoglobin precursor
44 (P01009) Alpha-1-antitrypsin precursor (Alpha-1 protease inhibitor) (Alpha-1-antiproteinase) (PRO06
45 (P01871) Ig mu chain C region
46 (P09871) Complement C1s subcomponent precursor (EC 3.4.21.42) (C1 esterase)
47 (Q8TE73) Ciliary dynein heavy chain 5 (Axonemal beta dynein heavy chain 5) (HL1)
48 (P02766) Transthyretin precursor (Prealbumin) (TBPA) (TTR) (ATTR)
49 (P07360) Complement component C8 gamma chain precursor
50 (P07358) Complement component C8 beta chain precursor (Complement component 8 beta subunit)
51 (P29622) Kallistatin precursor (Kallikrein inhibitor) (Protease inhibitor 4)
52 (P02649) Apolipoprotein E precursor (Apo-E)
53 (P39194) Alu subfamily SQ sequence contamination warning entry
54 (O95983) Methyl-CpG-binding domain protein 3 (Methyl-CpG binding protein MBD3)

55 (P35858) Insulin-like growth factor binding protein complex acid labile chain precursor (ALS)
56 (P02656) Apolipoprotein C-III precursor (Apo-CIII) (ApoC-III)
57 (P07996) Thrombospondin-1 precursor
58 (Q9NRC6) Spectrin beta chain, brain 4 (Spectrin, non-erythroid beta chain 4) (Beta-V spectrin) (BSP
59 (P39189) Alu subfamily SB sequence contamination warning entry
60 (P35442) Thrombospondin-2 precursor
61 (P15169) Carboxypeptidase N catalytic chain precursor (EC 3.4.17.3) (Lysine carboxypeptidase) (Argi
62 (Q9Y490) Talin 1
63 (Q08170) Splicing factor, arginine/serine-rich 4 (Pre-mRNA splicing factor SRP75) (SRP001LB)
64 (P10909) Clusterin precursor (Complement-associated protein SP-40,40) (Complement cytolysis inhibit
65 (P06753) Tropomyosin alpha 3 chain (Tropomyosin 3) (Tropomyosin gamma) (hTM5)
66 (Q99661) Kinesin-like protein KIF2C (Mitotic centromere-associated kinesin) (MCAK) (Kinesin-like pr
67 (Q9Y6E0) Serine/threonine-protein kinase 24 (EC 2.7.1.37) (STE20-like kinase MST3) (MST-3) (Mammali
68 (P39191) Alu subfamily SB2 sequence contamination warning entry
69 (O43781) Dual-specificity tyrosine-phosphorylation regulated kinase 3 (EC 2.7.1.-)
70 (O95479) GDH/6PGL endoplasmic bifunctional protein precursor [Includes: Glucose 1-dehydrogenase (EC
71 (P53396) ATP-citrate synthase (EC 2.3.3.8) (ATP-citrate (pro-S-)-lyase) (Citrate cleavage enzyme)
72 (P12111) Collagen alpha 3(VI) chain precursor
73 (O60237) Protein phosphatase 1 regulatory subunit 12B (Myosin phosphatase targeting subunit 2) (Myo
74 (Q96F44) Tripartite motif protein 11 (EC 6.3.2.-) (RING finger protein 92) (BIA1 protein)
75 (Q96T23) Hepatitis B virus x associated protein (HBV pX associated protein 8) (Remodeling and spaci
76 (O94973) Adapter-related protein complex 2 alpha 2 subunit (Alpha-adaptin C) (Adaptor protein compl
77 (P02647) Apolipoprotein A-I precursor (Apo-AI) (ApoA-I)
78 (O60741) Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 1 (Brain cycl
79 (P07205) Phosphoglycerate kinase, testis specific (EC 2.7.2.3)
80 (Q8IYF1) RNA polymerase II transcription factor SIII subunit A2 (Elongin A2) (EloA2) (Transcription
81 (O75874) Isocitrate dehydrogenase [NADP] cytoplasmic (EC 1.1.1.42) (Oxalosuccinate decarboxylase) (
82 (P06400) Retinoblastoma-associated protein (PP110) (P105-RB) (RB)
83 (Q9Y224) Protein C14orf166 (CGI-99)
84 (Q8IZT6) Abnormal spindle-like microcephaly-associated protein (Abnormal spindle protein homolog) (
85 (Q8WZ75) Roundabout homolog 4 precursor (Magic roundabout) (UNQ421/PRO3674)
86 (Q9NVW2) RING finger protein 12 (LIM domain interacting RING finger protein) (RING finger LIM domai
87 (Q15848) Adiponectin precursor (Adipocyte, C1q and collagen domain containing protein) (30 kDa adip
88 (O95425) Supervillin (Archvillin) (p205/p250)
89 (O14757) Serine/threonine-protein kinase Chk1 (EC 2.7.1.37)
90 (Q14152) Eukaryotic translation initiation factor 3 subunit 10 (eIF-3 theta) (eIF3 p167) (eIF3 p180
91 (Q9HCS7) XPA-binding protein 2 (HCNP protein) (PP3898)
92 (O75694) Nuclear pore complex protein Nup155 (Nucleoporin Nup155) (155 kDa nucleoporin)
93 (Q8TEU7) Rap guanine nucleotide exchange factor 6 (PDZ domain containing guanine nucleotide exchang
94 (Q96JA1) Leucine-rich repeats and immunoglobulin-like domains protein 1 precursor (LIG-1)
95 (Q9UHF7) Zinc finger transcription factor Trps1 (Zinc finger protein GC79) (Tricho-rhino-phalangeal
96 (P11532) Dystrophin
97 (P00451) Coagulation factor VIII precursor (Procoagulant component) (Antihemophilic factor) (AHF)
98 (Q03468) DNA excision repair protein ERCC-6 (Cockayne syndrome protein CSB)
99 (Q9NYQ6) Cadherin EGF LAG seven-pass G-type receptor 1 precursor (Flamingo homolog 2) (hFmi2)
100 (Q9UJV3) Midline 2 protein (Midline defect 2) (Tripartite motif protein 1) (Midin 2) (RING finger p
101 (O95782) Adapter-related protein complex 2 alpha 1 subunit (Alpha-adaptin A) (Adaptor protein compl
102 (P11021) 78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin heavy chain binding pr
103 (Q70EK9) Ubiquitin carboxyl-terminal hydrolase 51 (EC 3.1.2.15) (Ubiquitin thiolesterase 51) (Ubiqu
104 (Q96KN2) Glutamate carboxypeptidase-like protein 2 precursor (CNDP dipeptidase 1)
105 (Q93073) Protein KIAA0256
106 (Q9NQ55) Suppressor of SWI4 1 homolog (Ssf-1) (Peter Pan homolog)
107 (Q99728) BRCA1-associated RING domain protein 1 (BARD-1)
108 (Q14571) Inositol 1,4,5-trisphosphate receptor type 2 (Type 2 inositol 1,4,5-trisphosphate receptor
109 (Q9WJR5) HERV-K_19q12 provirus ancestral Pol protein (HERV-K(C19) Pol protein) [Includes: Reverse t
110 (P14625) Endoplasmin precursor (94 kDa glucose-regulated protein) (GRP94) (gp96 homolog) (Tumor rej
111 (O75533) Splicing factor 3B subunit 1 (Spliceosome associated protein 155) (SAP 155) (SF3b155) (Pre
112 (Q13029) PR-domain zinc finger protein 2 (Retinoblastoma protein-interacting zinc-finger protein) (
113 (O75417) DNA polymerase theta (EC 2.7.7.7) (DNA polymerase eta)
114 (Q9P2R3) Ankyrin repeat and FYVE domain protein 1 (Ankyrin repeats hooked to a zinc finger motif)
115 (P23229) Integrin alpha-6 precursor (VLA-6) (CD49f)
116 (Q13315) Serine-protein kinase ATM (EC 2.7.1.37) (Ataxia telangiectasia mutated) (A-T, mutated)
117 (Q9BUG6) Zinc finger and SCAN domain containing protein 5 (Zinc finger protein 495)
118 (Q03001) Bullous pemphigoid antigen 1 isoforms 1/2/3/4/5/8 (230 kDa bullous pemphigoid antigen) (BP

119 (O95970) Leucine-rich glioma-inactivated protein 1 precursor (Epitempin 1) (UNQ775/PRO1569)
120 (O95347) Structural maintenance of chromosome 2-like 1 protein (Chromosome-associated protein E) (h
121 (P05156) Complement factor I precursor (EC 3.4.21.45) (C3B/C4B inactivator)
122 (Q99973) Telomerase protein component 1 (Telomerase-associated protein 1) (Telomerase protein 1) (p
123 (Q9NSA2) Potassium voltage-gated channel subfamily D member 1 (Voltage-gated potassium channel subu
124 (P13535) Myosin heavy chain, skeletal muscle, perinatal (MyHC-perinatal)
125 (Q9NZV8) Potassium voltage-gated channel subfamily D member 2 (Voltage-gated potassium channel subu
126 (Q14289) Protein tyrosine kinase 2 beta (EC 2.7.1.112) (Focal adhesion kinase 2) (FADK 2) (Proline-
127 (Q9UQ05) Potassium voltage-gated channel subfamily H member 4 (Voltage-gated potassium channel subu
128 (Q9UGM5) Fetuin-B precursor (IRL685) (16G2)
129 (P42681) Tyrosine-protein kinase TXK (EC 2.7.1.112)
130 (O60469) Down syndrome cell adhesion molecule precursor (CHD2)
131 (P49895) Type I iodothyronine deiodinase (EC 1.97.1.10) (Type-I 5'deiodinase) (DIOI) (Type 1 DI) (5
132 (P02042) Hemoglobin delta chain
133 (Q8WWQ8) Stabilin 2 precursor (FEEL-2 protein) (Fasciclin EGF-like laminin-type EGF-like and link d
134 (Q96L93) Kinesin-like motor protein C20orf23 (Sorting nexin 23)
135 (Q9GZU5) Nyctalopin precursor
136 (P20848) Alpha-1-antitrypsin-related protein precursor
137 (Q9BXT8) RING finger protein 17
138 (O76039) Serine/threonine-protein kinase 9 (EC 2.7.1.37) (Cyclin-dependent kinase-like 5)
139 (P56715) Oxygen-regulated protein 1 (Retinitis pigmentosa RP1 protein) (Retinitis pigmentosa 1 prot
140 (P27169) Serum paraoxonase/arylesterase 1 (EC 3.1.1.2) (EC 3.1.8.1) (PON 1) (Serum aryldialkylphosp
141 (P20711) Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA decarboxylase) (DDC)
142 (Q86VI3) Ras GTPase-activating-like protein IQGAP3
143 (O15056) Synaptojanin 2 (EC 3.1.3.36) (Synaptic inositol-1,4,5-trisphosphate 5-phosphatase 2)
144 (Q13094) Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine
145 (O60563) Cyclin T1 (Cyclin T) (CycT1)
146 (Q9UIF8) Bromodomain adjacent to zinc finger domain 2B (hWALp4)
147 (Q14980) Nuclear mitotic apparatus protein 1 (NuMA protein) (SP-H antigen)
148 (P25054) Adenomatous polyposis coli protein (APC protein)
149 (P39190) Alu subfamily SB1 sequence contamination warning entry
150 (Q96RT6) Protein cTAGE-2
151 (Q15431) Synaptonemal complex protein 1 (SCP-1 protein)
152 (Q14683) Structural maintenance of chromosome 1-like 1 protein (SMC1alpha protein) (SB1.8/DXS423E p
153 (Q12923) Tyrosine-protein phosphatase, non-receptor type 13 (EC 3.1.3.48) (Protein-tyrosine phospha
154 (P20138) Myeloid cell surface antigen CD33 precursor (gp67) (Siglec-3)
155 (Q8WXH0) Nesprin 2 (Nuclear envelope spectrin repeat protein 2) (Syne-2) (Synaptic nuclear envelope
156 (Q9Y6N7) Roundabout homolog 1 precursor (H-Robo-1) (Deleted in U twenty twenty)
157 (P01861) Ig gamma-4 chain C region
158 (P02747) Complement C1q subcomponent, C chain precursor
159 (Q9Y4G6) Talin 2
160 (P43627) Killer cell immunoglobulin-like receptor 2DL2 precursor (MHC class I NK cell receptor) (Na
161 (P42566) Epidermal growth factor receptor substrate 15 (Protein Eps15) (AF-1p protein)
162 (Q9ULB1) Neurexin 1-alpha precursor (Neurexin I-alpha)
163 (Q9ULJ7) Hypothetical protein KIAA1223 (Fragment)
164 (Q9NQS7) Inner centromere protein
165 (Q15078) Cyclin-dependent kinase 5 activator 1 precursor (CDK5 activator 1) (Cyclin-dependent kinas
166 (Q01082) Spectrin beta chain, brain 1 (Spectrin, non-erythroid beta chain 1) (Beta-II spectrin) (Fo
167 (P10721) Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR) (Proto-oncogene tyro
168 (Q9Y4K1) Absent in melanoma 1 protein
169 (Q9UPN3) Microtubule-actin crosslinking factor 1, isoforms 1/2/3/5 (Actin cross-linking family prot
170 (Q8WXK8) Bullous pemphigoid antigen 1, isoform 7 (Bullous pemphigoid antigen) (BPA) (Hemidesmosomal
171 (Q06413) Myocyte-specific enhancer factor 2C
172 (P36955) Pigment epithelium-derived factor precursor (PEDF) (EPC-1)
173 (P36980) Complement factor H-related protein 2 precursor (FHR-2) (H factor-like protein 2) (H facto
174 (Q14896) Myosin-binding protein C, cardiac-type (Cardiac MyBP-C) (C-protein, cardiac muscle isoform
175 (Q9P2F8) Signal-induced proliferation-associated 1 like protein 2 (Fragment)
176 (Q9HCS5) Band 4.1-like protein 4A (NBL4 protein)
177 (Q9UGU5) High-mobility group protein 2-like 1 (HMGBCG protein)
178 (Q14573) Inositol 1,4,5-trisphosphate receptor type 3 (Type 3 inositol 1,4,5-trisphosphate receptor
179 (P13647) Keratin, type II cytoskeletal 5 (Cytokeratin 5) (K5) (CK 5) (58 kDa cytokeratin)
180 (Q8WXE1) ATR-interacting protein (ATM and Rad3 related interacting protein)
181 (Q8IZU2) WD-repeat protein 17
182 (O14910) LIN-7 homolog A (LIN-7A) (hLin-7) (Mammalian LIN-seven protein 1) (MALS-1) (Vertebrate LIN

183 (Q96I25) Splicing factor 45 (45kDa splicing factor) (RNA binding motif protein 17)
184 (O60229) Huntingtin-associated protein-interacting protein (Duo protein)
185 (Q9BVV6) Protein KIAA0586
186 (Q99828) Calcium and integrin-binding protein 1 (Calmyrin) (DNA-PKcs interacting protein) (Kinase i
187 (O75509) Tumor necrosis factor receptor superfamily member 21 precursor (TNFR-related death recepto
188 (O75376) Nuclear receptor corepressor 1 (N-CoR1) (N-CoR)
189 (P82279) Crumbs protein homolog 1 precursor
190 (P98164) Low-density lipoprotein receptor-related protein 2 precursor (Megalin) (Glycoprotein 330)
191 (Q8NHQ1) Centrosomal protein of 70 kDa (Cep70 protein) (p10-binding protein)
192 (Q9P2G1) Ankyrin repeat and IBR domain containing protein 1 (Fragment)
193 (Q9Y2I7) FYVE finger-containing phosphoinositide kinase (EC 2.7.1.68) (1-phosphatidylinositol-4-pho
194 (P12838) Neutrophil defensin 4 precursor (HNP-4) (HP-4) (Defensin, alpha 4)
195 (O75132) Zinc finger BED domain containing protein 4
196 (O14522) Receptor-type tyrosine-protein phosphatase T precursor (EC 3.1.3.48) (R-PTP-T) (RPTP-rho)
197 (Q9H0D6) 5'-3' exoribonuclease 2 (EC 3.1.11.-) (DHM1-like protein) (DHP protein)
198 (Q9Y2L1) Exosome complex exonuclease RRP44 (EC 3.1.13.-) (Ribosomal RNA processing protein 44) (DIS
199 (P55201) Peregrin (Bromodomain and PHD finger-containing protein 1) (BR140 protein)
200 (Q9UBZ9) DNA repair protein REV1 (EC 2.7.7.-) (Rev1-like terminal deoxycytidyl transferase) (Alpha
201 (Q8IZP9) G-protein coupled receptor 64 precursor (Epididymis-specific protein 6) (He6 receptor)
202 (Q9Y6D5) Brefeldin A-inhibited guanine nucleotide-exchange protein 2 (Brefeldin A-inhibited GEP 2)
203 (O75882) Attractin precursor (Mahogany homolog) (DPPT-L)
204 (O94782) Ubiquitin carboxyl-terminal hydrolase 1 (EC 3.1.2.15) (Ubiquitin thiolesterase 1) (Ubiquit
205 (O75095) Multiple EGF-like-domain protein 3 (Multiple epidermal growth factor-like domains 6) (Frag
206 (Q16787) Laminin alpha-3 chain precursor (Epiligrin 170 kDa subunit) (E170) (Nicein alpha subunit)
207 (Q9H4A3) Serine/threonine-protein kinase WNK1 (EC 2.7.1.37) (Protein kinase with no lysine 1) (Prot
208 (O95104) Splicing factor, arginine/serine-rich 15 (CTD-binding SR-like protein RA4)
209 (O95602) DNA-directed RNA polymerase I largest subunit (EC 2.7.7.6) (RNA polymerase I 194 kDa subun
210 (P07602) Proactivator polypeptide precursor [Contains: Saposin A (Protein A); Saposin B (Sphingolip
211 (Q9H3H1) tRNA isopentenyltransferase, mitochondrial precursor (EC 2.5.1.8) (Isopentenyl-diphosphate
212 (Q9C0H6) Kelch-like protein 4
213 (O14764) Gamma-aminobutyric-acid receptor delta subunit precursor (GABA(A) receptor)
214 (Q92692) Poliovirus receptor related protein 2 precursor (Herpes virus entry mediator B) (HveB) (Ne
215 (O94808) Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] 2 (EC 2.6.1.16) (Hexoseph
216 (Q13415) Origin recognition complex subunit 1 (Replication control protein 1)
217 (Q13105) Zinc finger and BTB domain containing protein 17 (Zinc finger protein 151) (Myc-interactin
218 (P67809) Nuclease sensitive element binding protein 1 (Y-box binding protein-1) (Y-box transcriptio
219 (Q14493) Histone RNA hairpin-binding protein (Histone stem-loop binding protein)
220 (P14923) Junction plakoglobin (Desmoplakin III)
221 (P50851) Lipopolysaccharide-responsive and beige-like anchor protein (CDC4-like protein) (Beige-lik
222 (Q96P65) Orexigenic neuropeptide QRFP receptor (G-protein coupled receptor 103) (SP9155) (AQ27)
223 (Q96PK2) Microtubule-actin crosslinking factor 1, isoform 4
224 (Q9UDV7) Zinc finger protein 282 (HTLV-I U5RE binding protein 1) (HUB-1)
225 (Q9Y5Y9) Sodium channel protein type X alpha subunit (Voltage-gated sodium channel alpha subunit Na
226 (Q9C040) Tripartite motif protein 2 (RING finger protein 86)
227 (P08697) Alpha-2-antiplasmin precursor (Alpha-2-plasmin inhibitor) (Alpha-2-PI) (Alpha-2-AP)
228 (P02787) Serotransferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal binding globulin) (PR
229 (P02786) Transferrin receptor protein 1 (TfR1) (TR) (TfR) (Trfr) (CD71 antigen) (T9) (p90)
230 (Q96RD0) Olfactory receptor 8B2
231 (O00625) Pirin
232 (O75161) Nephrocystin 4 (Nephroretinin)
233 (Q9UK17) Potassium voltage-gated channel subfamily D member 3 (Voltage-gated potassium channel subu
234 (Q02224) Centromeric protein E (CENP-E protein)
235 (Q9NQT4) Exosome complex exonuclease RRP46 (EC 3.1.13.-) (Ribosomal RNA processing protein 46) (Exo
236 (Q00059) Transcription factor A, mitochondrial precursor (mtTFA) (Mitochondrial transcription facto
237 (Q9Y2T1) Axin-2 (Axis inhibition protein 2) (Conductin) (Axin-like protein) (Axil)
238 (Q96DT6) Cysteine protease APG4C (EC 3.4.22.-) (Autophagy 4 homolog C) (Autophagin-3) (Autophagy-re
239 (Q9Y625) Glypican-6 precursor (UNQ369/PRO705)
240 (O60333) Kinesin-like protein KIF1B (Klp)
241 (Q14790) Caspase-8 precursor (EC 3.4.22.-) (CASP-8) (ICE-like apoptotic protease 5) (MORT1-associat
242 (Q9NZM3) Intersectin 2 (SH3 domain-containing protein 1B) (SH3P18) (SH3P18-like WASP associated pro
243 (Q92630) Dual-specificity tyrosine-phosphorylation regulated kinase 2 (EC 2.7.1.112) (EC 2.7.1.37)
244 (Q9P2L0) WD-repeat protein 35
245 (O15397) Importin 8 (Imp8) (Ran-binding protein 8) (RanBP8)
246 (Q14031) Collagen alpha 6(IV) chain precursor

247 (Q8NF91) Nesprin 1 (Nuclear envelope spectrin repeat protein 1) (Synaptic nuclear envelope protein
248 (O95171) Sciellin
249 (Q9NZJ4) Sacsin
250 (P39188) Alu subfamily J sequence contamination warning entry
251 (Q9Y4X5) Ariadne-1 protein homolog (ARI-1) (Ubiquitin-conjugating enzyme E2-binding protein 1) (Ubc
252 (Q16288) NT-3 growth factor receptor precursor (EC 2.7.1.112) (Neurotrophic tyrosine kinase recepto
253 (O75116) Rho-associated protein kinase 2 (EC 2.7.1.37) (Rho-associated, coiled-coil containing prot
254 (Q9C0C2) 182 kDa tankyrase 1-binding protein
255 (O60503) Adenylate cyclase, type IX (EC 4.6.1.1) (ATP pyrophosphate-lyase 9) (Adenylyl cyclase 9)
256 (Q8TF05) Serine/threonine phosphatase 4 regulatory subunit 1
257 (Q9BV73) Centrosomal protein 2 (Centrosomal Nek2-associated protein 1) (C-NAP1) (Centrosome protein
258 (O95402) Cofactor required for Sp1 transcriptional activation subunit 7 (Transcriptional coactivato
259 (Q9Y264) Angiopoietin-4 precursor (ANG-4) (ANG-3)
260 (P22352) Plasma glutathione peroxidase precursor (EC 1.11.1.9) (GSHPx-P) (Extracellular glutathione
261 (Q00839) Heterogenous nuclear ribonucleoprotein U (hnRNP U) (Scaffold attachment factor A) (SAF-A)
262 (P59910) Testis spermatocyte apoptosis-related gene 6 protein (Testis and spermatogenesis cell rela
263 (Q9HCK0) Zinc finger and BTB domain containing protein 26 (Zinc finger protein 481) (Zinc finger pr
264 (P61129) Zinc finger CCCH type domain containing protein 6
265 (Q13227) G protein pathway suppressor 2 (GPS2 protein)
266 (P63135) HERV-K_1q22 provirus ancestral Pol protein (HERV-K102 Pol protein) (HERV-K(III) Pol protei
267 (Q96ST3) Paired amphipathic helix protein Sin3a
268 (Q9UQM7) Calcium/calmodulin-dependent protein kinase type II alpha chain (EC 2.7.1.123) (CaM-kinase
269 (O15091) Hypothetical protein KIAA0391
270 (Q13620) Cullin homolog 4B (CUL-4B)
271 (Q8WYA0) Carnitine deficiency-associated protein expressed in ventricle 1 (CDV-1 protein)
272 (P28827) Receptor-type tyrosine-protein phosphatase mu precursor (EC 3.1.3.48) (Protein-tyrosine ph
273 (P16615) Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (EC 3.6.3.8) (Calcium pump 2) (SERCA2)
274 (P30414) NK-tumor recognition protein (Natural-killer cells cyclophilin-related protein) (NK-TR pro
275 (P26196) Probable ATP-dependent RNA helicase p54 (Oncogene RCK) (DEAD-box protein 6)
276 (Q9Y3L3) SH3-domain binding protein 1 (3BP-1)
277 (Q9Y4C0) Neurexin 3-alpha precursor (Neurexin III-alpha)
278 (Q9Y6D9) Mitotic spindle assembly checkpoint protein MAD1 (Mitotic arrest deficient-like protein 1)
279 (P50995) Annexin A11 (Annexin XI) (Calcyclin-associated annexin 50) (CAP-50) (56 kDa autoantigen)
280 (O43889) Cyclic-AMP responsive element binding protein 3 (Luman protein) (Transcription factor LZIP
281 (P06307) Cholecystokinins precursor (CCK) [Contains: Cholecystokinin 58 (CCK58); Cholecystokinin 39
282 (P48047) ATP synthase oligomycin sensitivity conferral protein, mitochondrial precursor (EC 3.6.3.1
283 (Q13506) NGFI-A binding protein 1 (EGR-1 binding protein 1) (Transcriptional regulatory protein p54
284 (Q9BZF2) Oxysterol binding protein-related protein 7 (OSBP-related protein 7) (ORP-7)
285 (Q14993) Collagen alpha 1(XIX) chain precursor (Collagen alpha 1(Y) chain)
286 (P05155) Plasma protease C1 inhibitor precursor (C1 Inh) (C1Inh)
287 (P35914) Hydroxymethylglutaryl-CoA lyase, mitochondrial precursor (EC 4.1.3.4) (HMG-CoA lyase) (HL)
288 (P54108) Cysteine-rich secretory protein-3 precursor (CRISP-3) (SGP28 protein)
289 (P17936) Insulin-like growth factor binding protein 3 precursor (IGFBP-3) (IBP-3) (IGF-binding prot
290 (P58340) Myeloid leukemia factor 1 (Myelodysplasia-myeloid leukemia factor 1)
291 (P50991) T-complex protein 1, delta subunit (TCP-1-delta) (CCT-delta) (Stimulator of TAR RNA bindin
292 (O00471) Exocyst complex component Sec10 (hSec10)
293 (Q14721) Potassium voltage-gated channel subfamily B member 1 (Voltage-gated potassium channel subu
294 (Q96F15) GTPase, IMAP family member 5 (Immunity-associated nucleotide 4-like 1 protein) (Immunity-a
295 (Q9UBX3) Mitochondrial dicarboxylate carrier
296 (P06241) Proto-oncogene tyrosine-protein kinase FYN (EC 2.7.1.112) (P59-FYN) (SYN) (SLK)
297 (Q8IWJ2) GRIP and coiled-coil domain-containing protein 2 (Golgi coiled coil protein GCC185) (CTCL
298 (O75791) GRB2-related adaptor protein 2 (GADS protein) (Growth factor receptor binding protein) (GR
299 (O15360) Fanconi anemia group A protein (FACA protein)
300 (P09668) Cathepsin H precursor (EC 3.4.22.16)
301 (O75113) Nedd4-binding protein 1 (N4BP1) (Fragment)
302 (P51826) LAF-4 protein (Lymphoid nuclear protein related to AF4)
303 (Q9UQE7) Structural maintenance of chromosome 3 (Chondroitin sulfate proteoglycan 6) (Chromosome-as
304 (Q04721) Neurogenic locus notch homolog protein 2 precursor (Notch 2) (hN2)
305 (P18206) Vinculin (Metavinculin)
306 (P24821) Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin) (GMEM) (JI) (Miotendinou
307 (Q8TC27) ADAM 32 precursor (A disintegrin and metalloprotease domain 32) (UNQ5982/PRO21340)
308 (P05543) Thyroxine-binding globulin precursor (T4-binding globulin)
309 (Q12789) General transcription factor 3C polypeptide 1 (Transcription factor IIIC-alpha subunit) (T
310 (Q08211) ATP-dependent RNA helicase A (Nuclear DNA helicase II) (NDH II) (DEAH-box protein 9)

311 (P57058) Hormonally up-regulated neu tumor-associated kinase (EC 2.7.1.37) (Serine/threonine-protei
312 (O14862) Interferon-inducible protein AIM2 (Absent in melanoma 2)
313 (Q9UMN6) Myeloid/lymphoid or mixed-lineage leukemia protein 4 (Trithorax homolog 2)
314 (Q02880) DNA topoisomerase II, beta isozyme (EC 5.99.1.3)
315 (P56192) Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase) (MetRS)
316 (P38935) DNA-binding protein SMUBP-2 (Immunoglobulin mu binding protein 2) (SMUBP-2) (Glial factor-
317 (P05164) Myeloperoxidase precursor (EC 1.11.1.7) (MPO)
318 (P31639) Sodium/glucose cotransporter 2 (Na(+)/glucose cotransporter 2) (Low affinity sodium-glucos
319 (Q86YR6) Ankyrin repeat domain protein 21 (POTE protein) (Prostate, ovary, testis expressed protein
320 (O15020) Spectrin beta chain, brain 2 (Spectrin, non-erythroid beta chain 2) (Beta-III spectrin)
321 (O95071) Ubiquitin--protein ligase EDD (EC 6.3.2.-) (Hyperplastic discs protein homolog) (hHYD) (Pr
322 (Q8IVF6) Ankyrin repeat domain protein 18A
323 (P12643) Bone morphogenetic protein 2 precursor (BMP-2) (BMP-2A)
324 (Q8TCX5) Rhophilin 1 (GTP-Rho binding protein 1)
325 (O00198) Activator of apoptosis harakiri (Neuronal death protein DP5) (BH3 interacting domain prote
326 (Q9UHA3) Probable ribosome biogenesis protein RLP24 (Ribosomal protein L24-like) (My024 protein)
327 (P33121) Long-chain-fatty-acid--CoA ligase 1 (EC 6.2.1.3) (Long-chain acyl-CoA synthetase 1) (LACS
328 (P49736) DNA replication licensing factor MCM2 (Minichromosome maintenance protein 2 homolog) (Nucl
329 (O60312) Potential phospholipid-transporting ATPase VA (EC 3.6.3.1) (ATPVA) (Aminophospholipid tran
330 (P17023) Zinc finger protein 19 (Zinc finger protein KOX12)
331 (O00555) Voltage-dependent P/Q-type calcium channel alpha-1A subunit (Voltage-gated calcium channel
332 (O60343) TBC1 domain family member 4
333 (Q9NUP9) LIN-7 homolog C (LIN-7C) (Mammalian LIN-seven protein 3) (MALS-3) (Vertebrate LIN 7 homolo
334 (Q9NVU0) DNA-directed RNA polymerases III 80 kDa polypeptide (EC 2.7.7.6) (RNA polymerase III subun
335 (O14795) Unc-13 homolog B (Munc13-2) (munc13)
336 (P48552) Nuclear factor RIP140 (Nuclear receptor interacting protein 1)
337 (Q9Y463) Dual-specificity tyrosine-phosphorylation regulated kinase 1B (EC 2.7.1.37) (EC 2.7.1.112)
338 (O43474) Kruppel-like factor 4 (Epithelial zinc-finger protein EZF) (Gut-enriched Krueppel-like fac
339 (Q9NQE7) Thymus-specific serine protease precursor (EC 3.4.-.-)
340 (P07814) Bifunctional aminoacyl-tRNA synthetase [Includes: Glutamyl-tRNA synthetase (EC 6.1.1.17) (
341 (Q6W2J9) BCoR protein (BCL-6 corepressor)
342 (P01213) Beta-neoendorphin-dynorphin precursor (Proenkephalin B) (Preprodynorphin) [Contains: Beta-
343 (O00622) CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61) (Insulin-like growth facto
344 (P26367) Paired box protein Pax-6 (Oculorhombin) (Aniridia, type II protein)
345 (Q9NQ66) 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta 1 (EC 3.1.4.11) (Phosphoino
346 (Q9P2K8) Eukaryotic translation initiation factor 2-alpha kinase 4 (EC 2.7.1.37) (GCN2-like protein
347 (P28289) Tropomodulin-1 (Erythrocyte tropomodulin) (E-Tmod)
348 (Q9P2R7) Succinyl-CoA ligase [ADP-forming] beta-chain, mitochondrial precursor (EC 6.2.1.5) (Succin
349 (Q15413) Ryanodine receptor 3 (Brain-type ryanodine receptor) (RyR3) (RYR-3) (Brain ryanodine recep
350 (O14983) Sarcoplasmic/endoplasmic reticulum calcium ATPase 1 (EC 3.6.3.8) (Calcium pump 1) (SERCA1)
351 (Q00653) Nuclear factor NF-kappa-B p100/p49 subunits (DNA-binding factor KBF2) (H2TF1) (Lymphocyte
352 (Q6UX41) Butyrophilin-like protein 8 precursor (UNQ702/PRO1347)
353 (O95819) Mitogen-activated protein kinase kinase kinase kinase 4 (EC 2.7.1.37) (MAPK/ERK kinase kin
354 (Q8TF76) Serine/threonine-protein kinase Haspin (EC 2.7.1.37) (Haploid germ cell-specific nuclear p
355 (Q9BZ29) Dedicator of cytokinesis protein 9 (Cdc42 guanine nucleotide exchange factor zizimin 1)
356 (O15270) Serine palmitoyltransferase 2 (EC 2.3.1.50) (Long chain base biosynthesis protein 2) (LCB
357 (Q92698) DNA repair and recombination protein RAD54-like (EC 3.6.1.-) (RAD54 homolog) (hRAD54) (hHR
358 (Q9Y2D4) Exocyst complex component Sec15B
359 (Q9UJF2) Ras GTPase-activating protein nGAP (RAS protein activator like 1)
360 (Q07283) Trichohyalin
361 (Q16513) Protein kinase N2 (EC 2.7.1.37) (Protein kinase C-like 2) (Protein-kinase C-related kinase
362 (O15294) UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit (EC 2.4.1
363 (O43927) Small inducible cytokine B13 precursor (CXCL13) (B lymphocyte chemoattractant) (CXC chemok
364 (Q15057) Centaurin beta 2 (Cnt-b2)
365 (P04275) Von Willebrand factor precursor (vWF) [Contains: Von Willebrand antigen II]
366 (P62244) 40S ribosomal protein S15a
367 (O15393) Transmembrane protease, serine 2 precursor (EC 3.4.21.-)
368 (Q9UPM8) Adapter-related protein complex 4 epsilon 1 subunit (Epsilon subunit of AP-4) (AP-4 adapte
369 (Q8NBK3) Sulfatase modifying factor 1 precursor (C-alpha-formylglycine-generating enzyme 1)
370 (P16157) Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R)
371 (Q9NR96) Toll-like receptor 9 precursor
372 (P15144) Aminopeptidase N (EC 3.4.11.2) (hAPN) (Alanyl aminopeptidase) (Microsomal aminopeptidase)
373 (Q8IY33) MICAL-like protein 2
374 (P35241) Radixin

375 (Q12816) Trophinin (MAGE-D3 antigen)
376 (P55084) Trifunctional enzyme beta subunit, mitochondrial precursor (TP-beta) (MSTP029) [Includes:
377 (Q9UL51) Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 2 (Brain cycl
378 (Q8NF50) Dedicator of cytokinesis protein 8 (Fragment)
379 (P14616) Insulin receptor-related protein precursor (EC 2.7.1.112) (IRR) (IR-related receptor)
380 (P20929) Nebulin
381 (P01112) Transforming protein p21/H-Ras-1 (c-H-ras)
382 (O15123) Angiopoietin-2 precursor (ANG-2)
383 (O14727) Apoptotic protease activating factor 1 (Apaf-1)
384 (P28715) DNA-repair protein complementing XP-G cells (Xeroderma pigmentosum group G complementing p
385 (Q9BXT6) Potential helicase Mov10l1 (EC 3.6.1.-) (Moloney leukemia virus 10-like protein 1) (MOV10-
386 (Q9UGQ3) Solute carrier family 2, facilitated glucose transporter, member 6 (Glucose transporter ty
387 (Q9H0N0) Ras-related protein Rab-6C (Rab6-like protein WTH3)
388 (P13639) Elongation factor 2 (EF-2)
389 (Q9Y5X9) Endothelial lipase precursor (EC 3.1.1.3) (Endothelial cell-derived lipase) (EDL) (EL)
390 (P27469) Putative lymphocyte G0/G1 switch protein 2
391 (Q9NSI8) SAM-domain protein SAMSN-1 (SAM domain, SH3 domain and nuclear localisation signals protei
392 (Q96A11) Galactose-3-O-sulfotransferase 3 (EC 2.8.2.-) (Gal3ST-3) (Galbeta1-3GalNAc 3'-sulfotransfe
393 (P47985) Ubiquinol-cytochrome c reductase iron-sulfur subunit, mitochondrial precursor (EC 1.10.2.2
394 (Q15746) Myosin light chain kinase, smooth muscle and non-muscle isozymes (EC 2.7.1.117) (MLCK) [Co
395 (Q8N8V4) Harmonin-interacting ankyrin-repeat containing protein (Harp)
396 (Q8N1I0) Dedicator of cytokinesis protein 4
397 (P08572) Collagen alpha 2(IV) chain precursor
398 (Q9UPT6) C-jun-amino-terminal kinase interacting protein 3 (JNK-interacting protein 3) (JIP-3) (JNK
399 (P34931) Heat shock 70 kDa protein 1L (Heat shock 70 kDa protein 1-like) (Heat shock 70 kDa protein
400 (Q96DT5) Ciliary dynein heavy chain 11 (Axonemal beta dynein heavy chain 11)
401 (O95935) T-box transcription factor TBX18 (T-box protein 18)
402 (Q75N90) Fibrillin 3 precursor
403 (Q9H8G1) Zinc finger protein 430
404 (Q9H5X1) Hypothetical UPF0195 protein FLJ22875
405 (P59531) Putative Taste receptor type 2 member 12 (T2R12) (Taste receptor type 2 member 26) (T2R26)
406 (Q8WVM8) Sec1 family domain containing protein 1 (Syntaxin binding protein 1-like 2) (Vesicle trans
407 (P11055) Myosin heavy chain, fast skeletal muscle, embryonic (Muscle embryonic myosin heavy chain)
408 (P39193) Alu subfamily SP sequence contamination warning entry
409 (Q92502) StAR-related lipid transfer protein 8 (StARD8) (START domain-containing protein 8)
410 (Q96SQ9) Cytochrome P450 2S1 (EC 1.14.14.1) (CYPIIS1) (UNQ891/PRO1906)
411 (Q9NUA8) Hypothetical zinc finger protein KIAA0478
412 (Q05586) Glutamate [NMDA] receptor subunit zeta 1 precursor (N-methyl-D-aspartate receptor subunit
413 (O15442) Adult brain protein 239 (239AB)
414 (O75071) Protein KIAA0494
415 (P07332) Proto-oncogene tyrosine-protein kinase Fes/Fps (EC 2.7.1.112) (C-Fes)
416 (Q9H9S5) Fukutin related protein (EC 2.-.-.-)
417 (P48995) Short transient receptor potential channel 1 (TrpC1) (TRP-1 protein)
418 (O95183) Vesicle-associated membrane protein 5 (VAMP-5) (Myobrevin) (HSPC191)
419 (Q9UM47) Neurogenic locus notch homolog protein 3 precursor (Notch 3)
420 (Q92805) Golgi autoantigen, golgin subfamily A member 1 (Golgin-97)
421 (Q06203) Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine phosphoribosylpyrophosph
422 (Q92616) GCN1-like protein 1 (HsGCN1)
423 (Q9NZR2) Low-density lipoprotein receptor-related protein 1B precursor (Low-density lipoprotein rec
424 (Q9HD67) Myosin X
425 (Q96M32) Putative adenylate kinase 7 (EC 2.7.4.3)
426 (P43220) Glucagon-like peptide 1 receptor precursor (GLP-1 receptor) (GLP-1-R) (GLP-1R)
427 (Q05996) Zona pellucida sperm-binding protein 2 precursor (Zona pellucida glycoprotein ZP2) (Zona p
428 (P47775) Probable G-protein coupled receptor 12
429 (O15084) Ankyrin repeat domain protein 28
430 (P49767) Vascular endothelial growth factor C precursor (VEGF-C) (Vascular endothelial growth facto
431 (P54803) Galactocerebrosidase precursor (EC 3.2.1.46) (GALCERase) (Galactosylceramidase) (Galactosy
432 (Q9NUE0) Zinc finger DHHC domain containing protein 18
433 (Q9P032) UPF0240 protein C6orf66 (HSPC125) (My013 protein)
434 (Q9UPV9) 106 kDa O-GlcNAc transferase-interacting protein
435 (Q14161) ARF GTPase-activating protein GIT2 (G protein-coupled receptor kinase-interactor 2) (GRK-i
436 (Q15417) Calponin-3 (Calponin, acidic isoform)
437 (Q13155) Multisynthetase complex auxiliary component p38 (JTV-1 protein) (PRO0992)
438 (Q14651) I-plastin (Intestine-specific plastin)

439 (O00482) Orphan nuclear receptor NR5A2 (Alpha-1-fetoprotein transcription factor) (Hepatocytic tran
440 (P54259) Atrophin-1 (Dentatorubral-pallidoluysian atrophy protein)
441 (Q96T68) Probable histone-lysine N-methyltransferase, H3 lysine-9 specific (EC 2.1.1.43) (Histone H
442 (O14511) Pro-neuregulin-2 precursor (Pro-NRG2) [Contains: Neuregulin-2 (NRG-2) (Neural-and thymus-d
443 (Q9BXL7) Caspase recruitment domain protein 11 (CARD-containing MAGUK protein 3) (Carma 1)
444 (Q14147) Probable ATP-dependent helicase DHX34 (DEAH-box protein 34)
445 (Q92610) Zinc finger protein 592
446 (Q96PY6) Serine/threonine-protein kinase Nek1 (EC 2.7.1.37) (NimA-related protein kinase 1) (NY-REN
447 (Q9P0W8) Spermatogenesis associated protein 7 (Spermatogenesis associated protein HSD3) (HSD-3.1)
448 (P51991) Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3)
449 (Q9H2K2) Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TNKS-2) (TRF1-interacting ankyrin-relate
450 (Q9P2M4) TBC1 domain family member 14
451 (Q9ULT8) HECT domain containing protein 1 (Fragment)
452 (O43399) Tumor protein D54 (hD54) (D52-like 2)
453 (P08908) 5-hydroxytryptamine 1A receptor (5-HT-1A) (Serotonin receptor 1A) (5-HT1A) (G-21)
454 (Q9H2U2) Inorganic pyrophosphatase 2, mitochondrial precursor (EC 3.6.1.1) (PPase 2) (Pyrophosphata
455 (Q9UL18) Eukaryotic translation initiation factor 2C 1 (eIF2C 1) (eIF-2C 1) (Putative RNA-binding p
456 (Q9UBM7) 7-dehydrocholesterol reductase (EC 1.3.1.21) (7-DHC reductase) (Sterol delta-7-reductase)
457 (Q13643) Skeletal muscle LIM-protein 2 (SLIM 2) (Four and a half LIM domains protein 3) (FHL-3)
458 (P98161) Polycystin 1 precursor (Autosomal dominant polycystic kidney disease protein 1)
459 (O94956) Solute carrier organic anion transporter family, member 2B1 (Solute carrier family 21, mem
460 (Q8NGA2) Olfactory receptor 7A2
461 (O94763) RNA polymerase II subunit 5-mediating protein (RPB5-mediating protein)
462 (Q99856) AT-rich interactive domain-containing protein 3A (ARID domain-containing protein 3A) (B-ce
463 (P61769) Beta-2-microglobulin precursor (HDCMA22P)
464 (Q9UQR0) Sex comb on midleg-like protein 2
465 (P21817) Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RyR1) (RYR-1) (Skeletal mu
466 (Q8TE04) Pantothenate kinase 1 (EC 2.7.1.33) (Pantothenic acid kinase 1) (hPanK1) (hPanK)
467 (Q9BZE4) Nucleolar GTP-binding protein 1 (Chronic renal failure gene protein) (GTP-binding protein
468 (P57071) PR-domain zinc finger protein 15 (Zinc finger protein 298)
469 (Q8N4C6) Ninein (hNinein)
470 (Q9UNA1) Rho-GTPase-activating protein 26 (Oligophrenin-1 like protein) (GTPase regulator associate
471 (P35916) Vascular endothelial growth factor receptor 3 precursor (EC 2.7.1.112) (VEGFR-3) (Tyrosine
472 (O43493) Trans-Golgi network integral membrane protein 2 precursor (Trans-Golgi network protein TGN
473 (Q92598) Heat-shock protein 105 kDa (Heat shock 110 kDa protein) (Antigen NY-CO-25)
474 (Q13459) Myosin IXb (Unconventional myosin-9b)
475 (Q8N3T1) Polypeptide N-acetylgalactosaminyltransferase-like protein 2 (EC 2.4.1.41) (Protein-UDP ac
476 (Q13573) Nuclear protein SkiP (Ski-interacting protein) (SNW1 protein) (Nuclear receptor coactivato
477 (P16118) 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 1 (6PF-2-K/Fru-2,6-P2ASE liver isozyme
478 (P13688) Carcinoembryonic antigen-related cell adhesion molecule 1 precursor (Biliary glycoprotein
479 (Q15772) Aortic preferentially expressed protein 1 (APEG-1)
480 (P83436) Conserved oligomeric Golgi complex component 7
481 (Q15648) Peroxisome proliferator-activated receptor binding protein (PBP) (PPAR binding protein) (T
482 (Q00722) 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta 2 (EC 3.1.4.11) (Phosphoino
483 (Q13435) Splicing factor 3B subunit 2 (Spliceosome associated protein 145) (SAP 145) (SF3b150) (Pre
484 (O15297) Protein phosphatase 2C delta isoform (EC 3.1.3.16) (PP2C-delta) (p53-induced protein phosp
485 (Q9H2T7) Ran-binding protein 17
486 (Q9C0D0) Phosphatase and actin regulator 1
487 (Q15596) Nuclear receptor coactivator 2 (NCoA-2) (Transcriptional intermediary factor 2)
488 (Q9Y252) RING finger protein 6 (RING-H2 protein)
489 (P42356) Phosphatidylinositol 4-kinase alpha (EC 2.7.1.67) (PI4-kinase) (PtdIns-4-kinase) (PI4K-alp
490 (Q8IUD2) ERC protein 1 (ELKS protein)
491 (P00367) Glutamate dehydrogenase 1, mitochondrial precursor (EC 1.4.1.3) (GDH)
492 (P49641) Alpha-mannosidase IIx (EC 3.2.1.114) (Mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase)
493 (Q9Y2X9) Zinc finger protein 281 (Zinc finger DNA binding protein 99) (Transcription factor ZBP-99)
494 (O95210) Genethonin 1 (GENX-3414)
495 (O75151) PHD finger protein 2 (GRC5)
496 (P78563) Double-stranded RNA-specific editase 1 (EC 3.5.-.-) (dsRNA adenosine deaminase) (RNA editi
497 (P62910) 60S ribosomal protein L32 (PP9932)
498 (P08621) U1 small nuclear ribonucleoprotein 70 kDa (U1 snRNP 70 kDa) (snRNP70) (U1-70K)
499 (Q9H2G9) Golgin 45 (Basic leucine zipper nuclear factor 1) (JEM-1) (p45 basic leucine-zipper nuclea
500 (Q99460) 26S proteasome non-ATPase regulatory subunit 1 (26S proteasome regulatory subunit RPN2) (2

Appendix 4

Draft of manuscript describing development of our proteomic methods

DRAFT ONLY

Serologic biomarker discovery for human breast cancer using proteomics approach-method development

Yansen Xiao[1], Eddie Nieves[1], Ruth Angeletti[1], Mindy Ginsberg[2], Joe Vogelman[3], Laurel Habel[4], Peiching Du[1], Y.G. Yeung[1], E. Richard Stanley[1], Thomas Rohan[2,5,6]

[1] Department of Developmental and Molecular Biology, Albert Einstein College of Medicine

[2] Department of Epidemiology and Population Health, Albert Einstein College of Medicine

[3] Orentreich Foundation for the Advancement of Science

[4] Kaiser Permanente Medical Care Program

[5] Correspondence: Dr. T. Rohan, Dept. of Epidemiology and Population Health, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461

[6] Supported by funding from the Department of Defense Breast Cancer Research Program (.W-81XWH-05-1-0298)

Abstract

To search for biomarkers predictive of breast cancer risk, sera from normal subjects and women who developed breast cancer within 5 years of sample collection were analyzed in parallel using proteomics approaches. The serum samples of each case and its matched control were analyzed in the same run. To provide the dynamic range required by proteomic analysis, the samples underwent multiple steps of fractionation. Initially, the sera were depleted of major abundant proteins by passing the samples through a 12-protein immunoaffinity column. The resulting depleted samples were subjected to centrifugation filtration to concentrate the proteins and to change the buffer for trypsin digestion. A pooled sample, created by mixing an aliquot of the serum samples from all included subjects, and included in each run, was subjected to the same procedures. After overnight digestion, the digests were labeled individually with one of the multiplex iTRAQ reagents. The labeled digests were then mixed as one sample and were separated into six fractions with strong cation exchange chromatography (SCX). Each of the 6 SCX fractions was further separated into 192 subfractions by capillary C18 reversed phase liquid chromatography coupled with robotic auto spotting onto 192-well MALDI (matrix assisted laser desorption ionization) plates. The MALDI plates were analyzed using the 4800? Proteomic Analyzer mass spectrometer. The top 25 ? parent ions acquired in reflector mode from each of the 192 spots were fragmented under CID (collision induced dissociation) conditions to generate peptide sequence data as well as quantification data. Finally, the obtained data were searched using the GPS program to reveal the identity of proteins and their relative quantity in normal subjects and in those who subsequently developed breast cancer.

Introduction

Breast cancer is the most commonly diagnosed solid tumor and the second leading cause of death from cancer among women in the United States (American Cancer Society. Cancer Facts and Figures: 2008: Atlanta. 2008. Prediction of those who are at risk is essential for prevention. Breast cancer tumorigenesis involves numerous events at the molecular level during its growth, invasion, and metastasis [2, 3, 4]. It has been hypothesized that these changes will be reflected eventually in the protein profile of serum, due to endocrinological modulation, immuno/inflammatory response, cell damage, and tissue reconstruction, etc. The number of proteins potentially involved is so large and their dynamic ranges vary so markedly that common biochemical methods cannot be used for protein profiling of breast cancer tumorigenesis. Fortunately, recent advances in proteomics with mass spectrometry provide an ideal tool for protein profiling in a given sample like serum [5,6] . The shotgun proteomics strategy, by enzymatically digesting the proteins in mixture followed by iTRAQ labeling and mass spectrometric detection of digested peptides, can be used to identify and quantify polypeptides and proteins. This procedure is now being used in biomarker studies of several diseases [7,8,9,10,].

Breast cancer biomarkers might be present in different types of body fluids such as nipple aspirate fluid, serum, and urine [13]. However, because of the simplicity of sample preparation, serum or plasma has been used most frequently for biomarker analysis [14, 15]. Of the serum proteins, 99% are highly abundant proteins such as albumin, immunoglobulins, haptoglobin, and lipoproteins, etc. These abundant proteins will certainly mask the detection of possible biomarkers that are in the lower end of the dynamic range. The most advanced method in proteomics analysis is first to remove/deplete the abundant proteins by immunoaffinity depletion chromatography [11,12]. Any variation in this process will cause considerable error in the detection of biomarkers. The subsequent steps of sample processing, such as strong cation exchange chromatography and

reversed phase liquid chromatography, are also sources of analytical biases if not established properly. Therefore, for breast cancer biomarker discovery, it is necessary to explore the reproducibility of each sample fractionation step to minimize experimental errors and biases. We describe here a method for efficiently identifying possible candidate proteins/polypeptides biomarkers in serum.

Materials and Methods

To explore possible candidates for breast cancer biomarkers, sera from normal subjects (controls) and those who developed breast cancer (cases) within 5 years of collection of the serum sample were selected from the cohort of women who were members of the Northern California Region Kaiser Permanente Medical Care Program and who underwent a multi-phasic health check-up (MHC) between 1985 and 1992.. A pooled serum sample was created by mixing aliquots of serum samples from all cases and controls selected for the study, and pooled samples were used for the work described here.

1. ***Subject selection*** This project will be conducted in two phases: a training phase and a test phase. Each phase will be conducted as a case-control study nested in the MHC cohort. The purpose of the training phase is to identify biomarkers that discriminate between women who developed invasive breast cancer within 5 years of having a serum sample collected and women who remained free of breast cancer for at least as long as their corresponding case. The purpose of the test phase is to test the biomarkers identified, in the training phase using an independent (validation) nested case-control set of serum samples. Cases are defined as white, postmenopausal women, aged 55 to 80 years with no history of breast cancer at the time of recruitment and whose breast cancer was identified by merging data from the multiphaseic cohort/serum repository databases with data from the Kaiser Permanente tumor registry. Controls are matched 1:1 to the corresponding case and had

not developed breast cancer by the data of diagnosis of the corresponding case. The controls were selected using risk-set sampling with replacement, with matching to the corresponding case on age (within one year) and date of serum collection (within one month). Furthermore, cases and controls were matched with respect to membership of Kaiser Permanente, that is, controls were required to have been members from one year prior to serum collection and to have been a member at the time of diagnosis of the corresponding case. Finally, controls were matched to cases on time of blood draw with respect to the last meal: 0-3 hours of 4-9 hours since last meal. Study identification numbers were assigned to the cases and controls. These numbers were linked to the Kaiser IDs, which allowed extraction of corresponding covariate information from the cohort database.

2. ***Serum collection and aliquotation*** Blood was drawn within 9 hours of the last meal. The pulling, testing (for dessication), and aliquoting of the serum specimens was completed at the Orentreich Foundation, where the serum repository is located. For each subject included in the study, 5 x 20 μl aliquots were made. Furthermore, a common serum pool was created by adding 50 μl from each study subject (cases and controls) to the pool. The pool serves as a common standard that will be included in each run. Therefore, each of the sixty strata included in this study has 5 sets of identical triplets. Each triplet contains a case, a control, and a pooled serum sample, each of which has a volume of 20 μl. The location of the case and control aliquot within each triplet was assigned randomly and the laboratory staff were blinded to the identity of the case and control samples. The specimens were shipped to and received at the Albert Einstein College of Medicine where they are held in storage at -80^{o}C until needed for analysis.

3. ***Proteomics analysis of serum*** The work flow of sample processing and proteomic analysis is shown in Figure 1. Details of each step are described in the following sections.

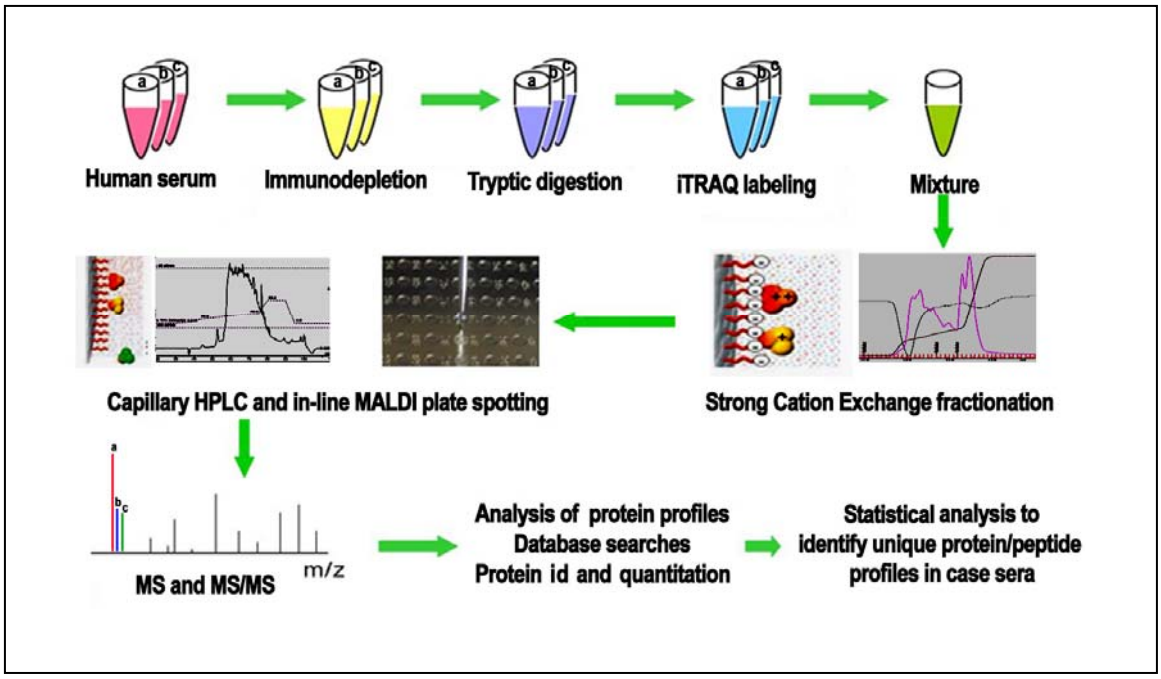


Fig. 1. Work flow of this study. Three sera from one stratum were depleted, digested, and iTRAQ labeled separately, followed by mixing, SCX fractionation, LC fractionation and spotting to a MALDI target. The final samples were analyzed with MALDI/TOF/TOF for protein identification and quantification.

3.1. **Immunodepletion of abundant proteins from serum** The samples were analyzed one triplet at a time. After thawing on ice, 80 μl of phosphate buffered saline (Sigma, St. Louis, MO) was added [need to clarify here – in this paper, are we going to describe work based on the pooled samples only?] to the 20 μl serum samples for the case, control, and pooled sample. The diluted serum preparation was loaded onto a 12-protein immunodepletion column (6.4 × 63 mm, Genway Biotech Inc, San Diego, CA) using the AKTA Purifier 10 system (GE Healthcare Bio-Sciences, Piscataway, NJ). These proteins include HAS, IgG, fibrinogen, transferrin, IgA, IgM, haptoglobin, a2-macroglobulin, a1-acid glycoprotein, a1-antitrypsin, Apo A-I and Apo A-II that collectively

comprise nearly 95% of total serum proteins by weight. Flow-through fractions of 3 ml in PBS were collected as monitored by UV absorbance at wave length 280nm. The column was regenerated by eluting with 5 column volumes of 100 mM glycine-HCl, pH 2.5, and neutralizing with 100 mM Na_2HPO_4-NaH_2PO_4, pH 8.0, followed by washing with PBS. The column was preserved in PBS with 0.01% sodium azide at 4°C after daily use. Depletion and elution profiles are shown in Figure 2 as an example.

3.2. **Concentration, buffer exchange, and trypsin digestion of depleted sera** The collected 3 ml depleted serum from each member of the triplet was concentrated using an Amicon Ultra centrifugal filter tube with 5 kDa Mr cutoff as described by the manufacturer (Millipore, Bedford, MA). Briefly, the devices were pre-incubated with 5% Triton X-100 for one hour on ice followed by rinsing with Milli-Q H_2O (Millipore, Bedford, MA). The 3 ml depleted serum was transferred to the filter tube and centrifugation was done at 3950 rpm at 8 ^oC for 50 min to reduce the volume to less than 80 μl. The spin-thru was discarded and 3 ml of 50 mM triethylammonium bicarbonate buffer (TEAB, Sigma-Aldrich, Saint Louis, MO), pH 8.5, was added to the concentrate to change the buffer for trypsin digestion. After gentle resuspension, the sample was centrifuged again as above. The last spin was repeated once to reduce the volume to less than 80μl and to reduce the salt concentration to less than 1 mM/???. The final concentrate was reconstituted to 100 μl of volume containing 20% acetonitrile, 50 mM TEAB. 5 μl of digestion solution containing 5 μg of sequencing grade TPCK treated trypsin (Promega, Madison, WI) freshly dissolved in 50 mM TEAB, was added to each serum sample and was incubated at 37 ^oC overnight at 400 rpm on a Thermomixer R device (Eppendorf North America, Westbury, NY).

3.3. **iTRAQ labeling and SCX fractions of digests** For relative quantification of the peptides among the three sera in a stratum, the digests were labeled in parallel with iTRAQ reagent 114, 115 or 116 (Applied Biosystem, Foster City, CA). In a stratum triplet, one of the serum samples was labeled with reagent 114, one with reagent 115, and one with reagent 116. The labeling reaction was performed according to the manufacturer's instructions with modifications. The three vials of 114, 115, and 116 were moved from -2??80 ^{o}C to room temperature 30 min before opening. Then 105μl of ethanol was added to each vial to dissolve the reagent. 40 μl digest from each sample was transferred to each corresponding iTRAQ vial. The 3 vials were mixed, spun briefly, and incubated at room temperature overnight. After incubation, the reaction was stopped by adding 420 μl H2O and incubating for 30 min at room temperature. For subsequent SCX fractionation, 466 μl of the labeled digest from each sample was mixed with SCX buffer A (10 mM potassium phosphate, 25% acetonitrile, pH 3.0) to a final volume of 5.7 ml. SCX chromatography was done on a PolySulfoethyl A SCX column (2.1 × 100 mm, 5 μm, 300Å; Poly LC Inc, Columbia, MD) using the AKTA Purifier 10 system. The mixture of the labeled digest was loaded to the SCX column and washed thoroughly with SCX buffer A. Elution of bound peptides was performed with SCX buffer B containing the same components as SCX buffer A except that 700 mM of KCl was added. The gradient of SCX buffer B is as described in the following steps: 0 to 10% B for 2 min, 10 to 25% B for 17 min, and 25 to 100% B for 5 min. The flow rate is 0.1 ml/min during elution. Elution profiles were monitored with UV at 214 mm. The eluent was collected into six 400 μl fractions and stored at 4 ^{o}C.

3.4. **C18 capillary reversed phase liquid chromatography (RPLC) and in-line MALDI target spotting** The volumes of the SCX fractions were reduced to about 80 μl in Speed Vac Concentrator (Savant Instrument Inc., Farmingdale, NY). They were then reconstituted to 100 μl containing 5%

acetonitrile, 0.1% trifluoroacetic acid (TFA) and transferred to autosampler vials (500 μl) for the Ultimat 3000 capillary HPLC system (Dionex, Sunnyvale, CA). The sample was first loaded to a μ-Precolumn (C18 PepMap 100, 5μm, 100A, 1.0 mm i.d. x 15 mm, Dionex, Sunnyvale, CA) in μ-Pickup injection mode with LC Solution A containing 95% H_2O, 5% acetontrile and 0.1% TFA at a rate of 40 μl/min with LoadingPump. After washing with Solution A, the precolumn was connected to a C18 capillary column (C18 PepMap100, 100A, 300 μm i.d. x 15 cm, Dionex, Sunnyvale, CA) and elution was started with MicroPump at the same time at a flow rate of 2 μl/min. Elution solution (Solution B) contains 5% H_2O, 95% acetontrile, and 0.1% TFA. During elution, the gradient of Solution B was set as the following: 5 to 12% for 10 min, 12 to 20% for 3 min, 20 to 40% for 27 min, 40 to 40% for 5 min, and 40 to 90% for 5 min. Elution profiles were monitored with UV at 214 nm. Eluent was connected to a Bai Probot MALDI spotting device (Dionex, Sunnyvale, CA), synchronized with MALDI matrix delivery, and spotted automatically to 192-well ABI 4700 Proteomics Analyzer plate insert. The matrix contained 6.2 mg/ml of alpha-cyano-4-hydroxycinnamic acid in 36% methanol, 56% acetonitrile, and 8% H2O (Agilent, New Castle, DE). Matrix and eluent were mixed in a ratio of 1:1. Each well collected 12.5 seconds of eluent-matrix mixture. Eluent from each SCX fraction was spotted to one plate insert.

3.5. **MALDI TOF MS/MS analysis of peptides** The samples spotted on the MALDI plate insert were analyzed using the 4800 Proteomics Analyzer (Applied Biosystems, Framingham, MA). Spot set template was set as LC-MALDI run. Mass accuracy was calibrated with the standard peptide mixture (des-Arg1-Bradykinin, 904.5 Da; Angiotensin, 1,296.7 Da; Glu-Fibrinopeptide B, 1,570.7 Da; ACTH1-17, 2.093.1 Da; ACTH18-39, 2,465.2 Da, and ACTH7-38, 3,657.9 Da. Applied Biosystems, Framingham, MA) loaded manually to the six calibration spots. For service scan in reflector mode, 1500 shots (50 shots/spectrum) were taken. The top 25 high intensity peaks were

selected for MS/MS analysis in CID (collision induced dissociation) mode. 1500 shots (100 shots/spectrum) were taken for each MS/MS scan.

3.6. **Data analysis and protein identification** MS/MS data from each stratum were searched for protein identity using the GPS protein search software [1].

Results

1. Efficiency of immunodepletion and trypsin digestion

The immunodepletion chromatography column used in this study is designed to remove the 12 ?change??most abundant serum proteins in 25 µl of serum. This column can be regenerated after each run for at least 100 times with only minor loss in its capacity. The general depletion profiles of three pooled samples, as monitored using UV absorbance at 280 nm, were nearly identical (Figure 2). This indicates that this step is reproducibile. In our samples, based on the measurement of the peak areas, about 75% proteins were removed from the serum in each depletion run.

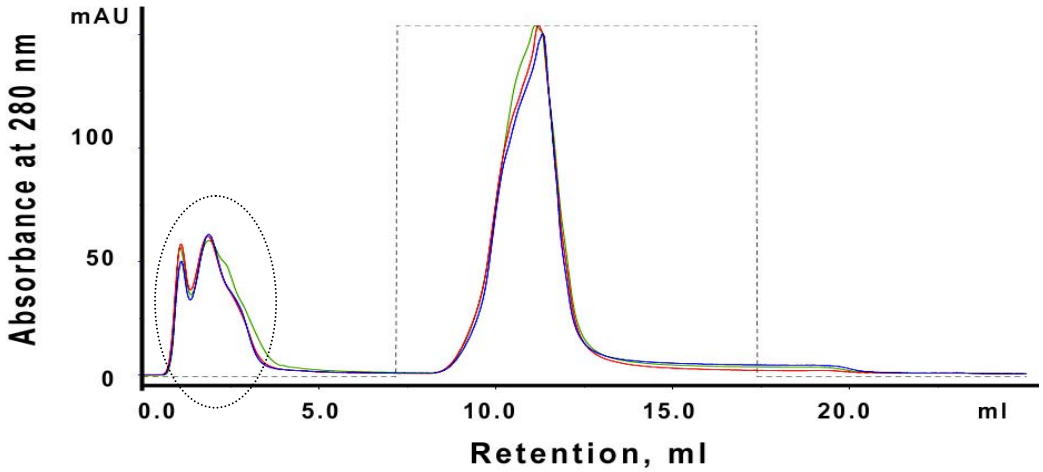


Figure 2. Immunoaffinity depletion of abundant serum proteins using an affinity column with 12 immobilized IgY antibodies. Depletion profiles of three pooled sera are overlaid to demonstrate the reproducibility of this procedure. The circled peaks are the fractions that are depleted of the 12 most abundant proteins and are saved for further processing – not clear. The second peaks are the abundant proteins that are removed from the sera and discarded.

The depleted sera, 3 ml each, were concentrated individually using centrifugal filtration tubes with 5-kDa molecular weight cut. After the first spin, the approximately 50 μl of concentrate was resuspended in 3 ml of 50 mM TEAB buffer to further reduce the salt concentration and to change buffer for subsequent trypsin digestion. This was repeated once. To evaluate the recovery of proteins after each procedure, equivalent amounts of protein from each step were taken for SDS-PAGE analysis. Results are shown in Figure 3. It is clear that there is no noticeable lose of protein during this procedure.

For trypsin digestion, the 50 μl concentrate was digested over night with 5 μg of proteomic grade trypsin in 50 mM TEAB containing 20% acetonitrile. The introduction of acetonitrile is to help the digestion of hydrophobic proteins as well as to make the medium compatible with the next labeling step [1]. The presence of acetonitrile in the digestion buffer has no adverse effect on the digestion as detected using SDS-PAGE gel (data not shown). Under the conditions used for trypsin digestion, the proteins are completely digested, as revealed by silver stain analysis of SDS-PAGE gel (Figure 3).

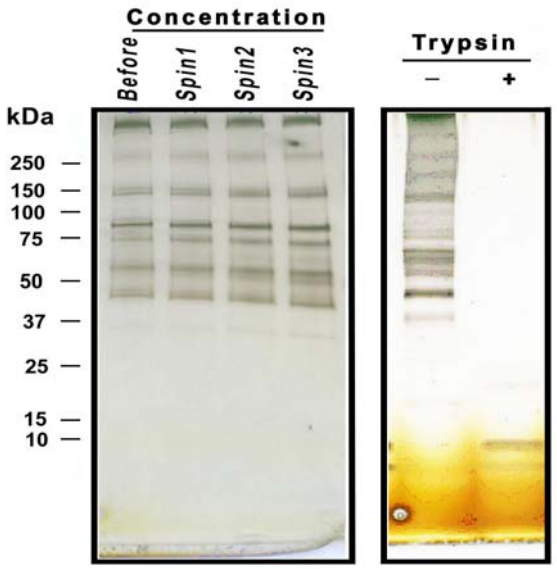


Fig. 3. Sample recovery after centrifugal filtration/concentration (Left) and efficiency of trypsin digestion. In the left gel, equal amounts of protein from the starting sample and those after each spin were loaded in to each well of an SDS-PAGE gel. The gel was stained with silver salt to reveal the protein bands. In the right gel, sample was taken before and after digestion

2. iTRAQ labeling

To ensure that the iTRAQ labeling is complete and different isotopes label equally, a standard peptide (bradykinin) was used to test the labeling efficiency before labeling the serum digests. Under the conditions used in this study, bradykinin was completely labeled. When analyzed in ms/ms mode, the ions representing 114, 115, and 116 are clearly detected and their intensities (peak areas) are similar. The small variation might be caused by sampling because the test was done ?using a few µl of media (Figure 4).

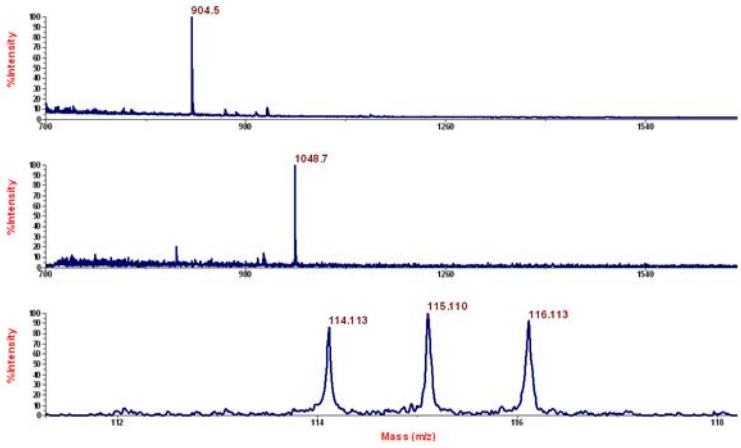


Fig. 4. Test of iTRAQ labeling efficiency. Bradykinin peptide was labeled with 114, 115, and 116 separately. After 30 min labeling, the reaction was stopped with H2O and was mixed at equal volume. MALDI-TOF was performed to examine the labeling efficiency. Upper spectrum: before labeling, unmodified bradykinin was detected at 904.5 Da. Mid spectrum: MALDI-TOF analysis of peptide mixture after labeling, indicating that all bradykinin was labeled into the peptide at 1048.7 Da. Lower spectrum: MALDI-TOF-TOF analysis of the labeled mixture, indicating that different isotopes label equally.

In the present study, the efficiency of iTRAQ labeling was evaluated as the percentage of labeled peptides. The 114, 115, or 116 iTRAQ reagent cluster areas as obtained from GPS search result data were compared with those of 117. Here the 117 areas are virtual iTRAQ reagent cluster areas since we did not use this isotope. The areas of 117 serve only as background noise. An iTRAQ peptide is defined as one that has a ratio of 114/117??, 115/117 or 116/117 greater than 10. The details of the calculation are listed in Table 1. As shown, in both sets (strata) of samples, the labeling efficiency is around 99%, which is acceptable.

Table 1. Efficiency of iTRAQ labeling. A peptide is defined as a iTRAQ labeled peptide when the ratio of the cluster areas of 114/117, 115/117 or 116/117 is greater than 10. The percentage of iTRAQ labeling is calculated by comparing the number of iTRAQ peptides with the total number of MS/MS scans.

	No. of iTRAQ peptides			Total MS/MS scans	Percentage of iTRAQ labeling		
	114	115	116		114	115	116
Stratum1	17621	17580	17633	17744	99.31	99.08	99.37
Stratum2	13446	13488	13457	13606	98.82	99.13	98.90

3. Strong Cation Exchange fractionation of tryptic digest

Although depleted of the top abundant proteins, the iTRAQ labeled tryptic digests contain still contain numerous peptides. Direct analysis of the digest using MALDI mass spectrometry is impossible due to an ion suppression effect. Therefore, further fractionation is mandatory. Here, strong cation exchange chromatography was used to separate the peptide mixture into several fractions according to the charges they bear. This was done on a PolySULFOETHYL Aspartamide SCX column.

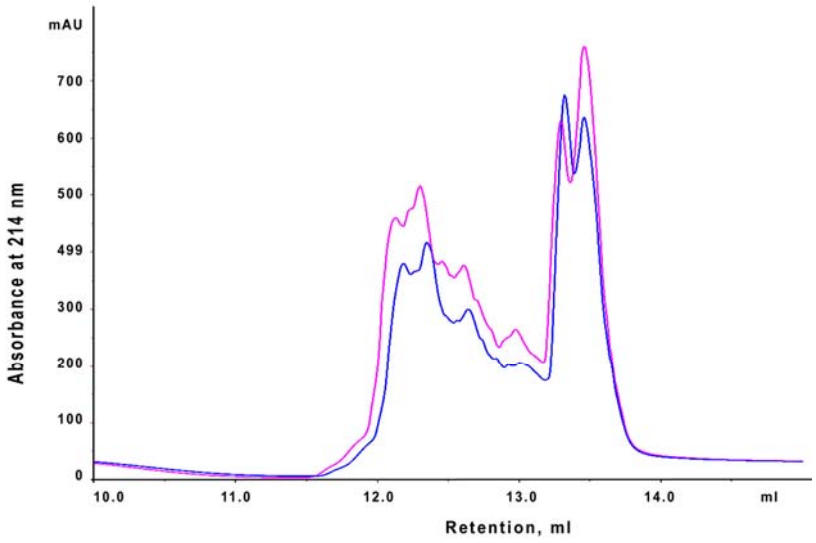


Fig.5. Strong Cation Exchange fractionation of tryptic digest.

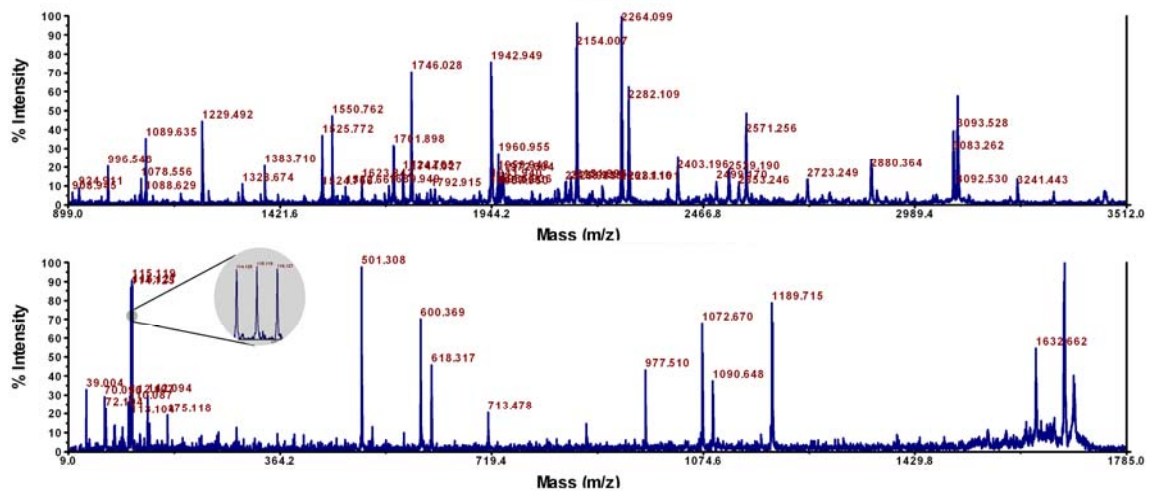


Fig. 6. Example of iTRAQ labeling of digest. The upper spectrum is acquired in reflectoron mode of one fraction

4. Capillary Reversed phase liquid chromatography of SCX fractionated samles

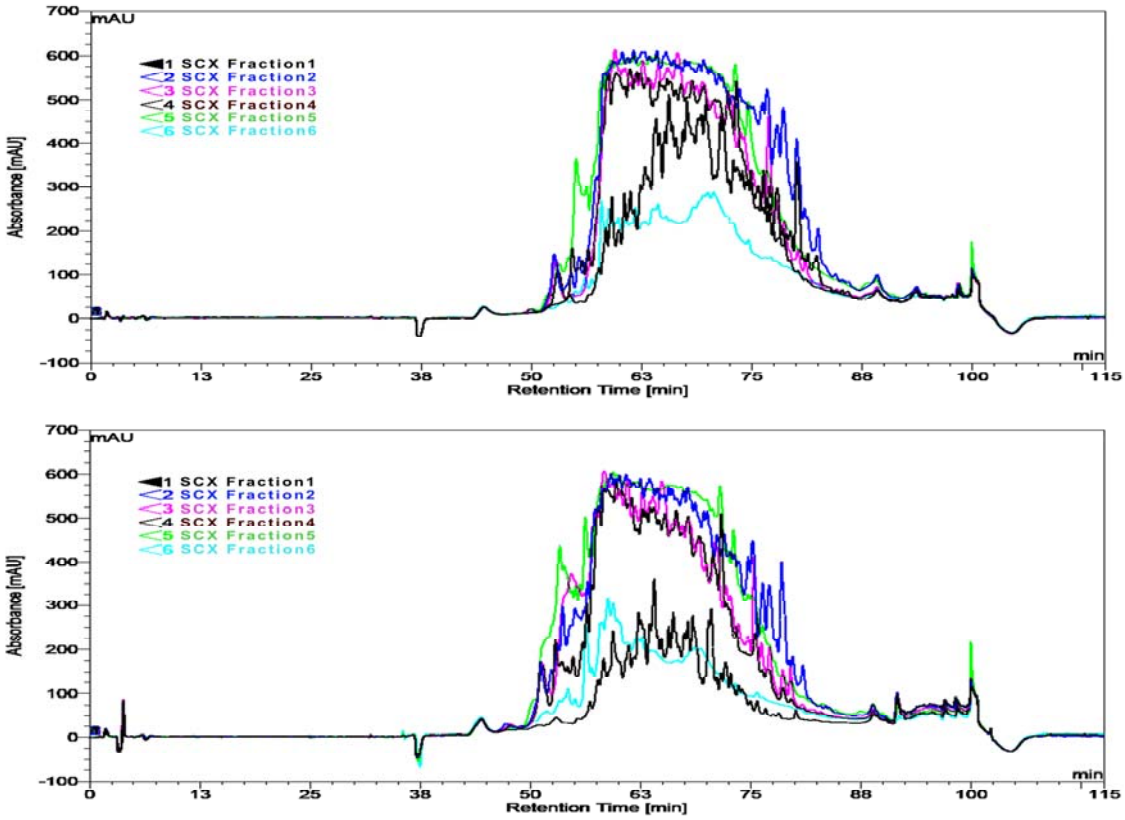


Fig. 5 Comparison of LC profiles of two repeat strata. Six combined SCX fractions from each stratum were separated using capillary RPLC and eluted with a step gradient of acetronitrile of up to 95%.

Fig. 7. An overlay of LC profiles of six SCX fractions from on e stratum

5. MALDI TOF/TOF analysis

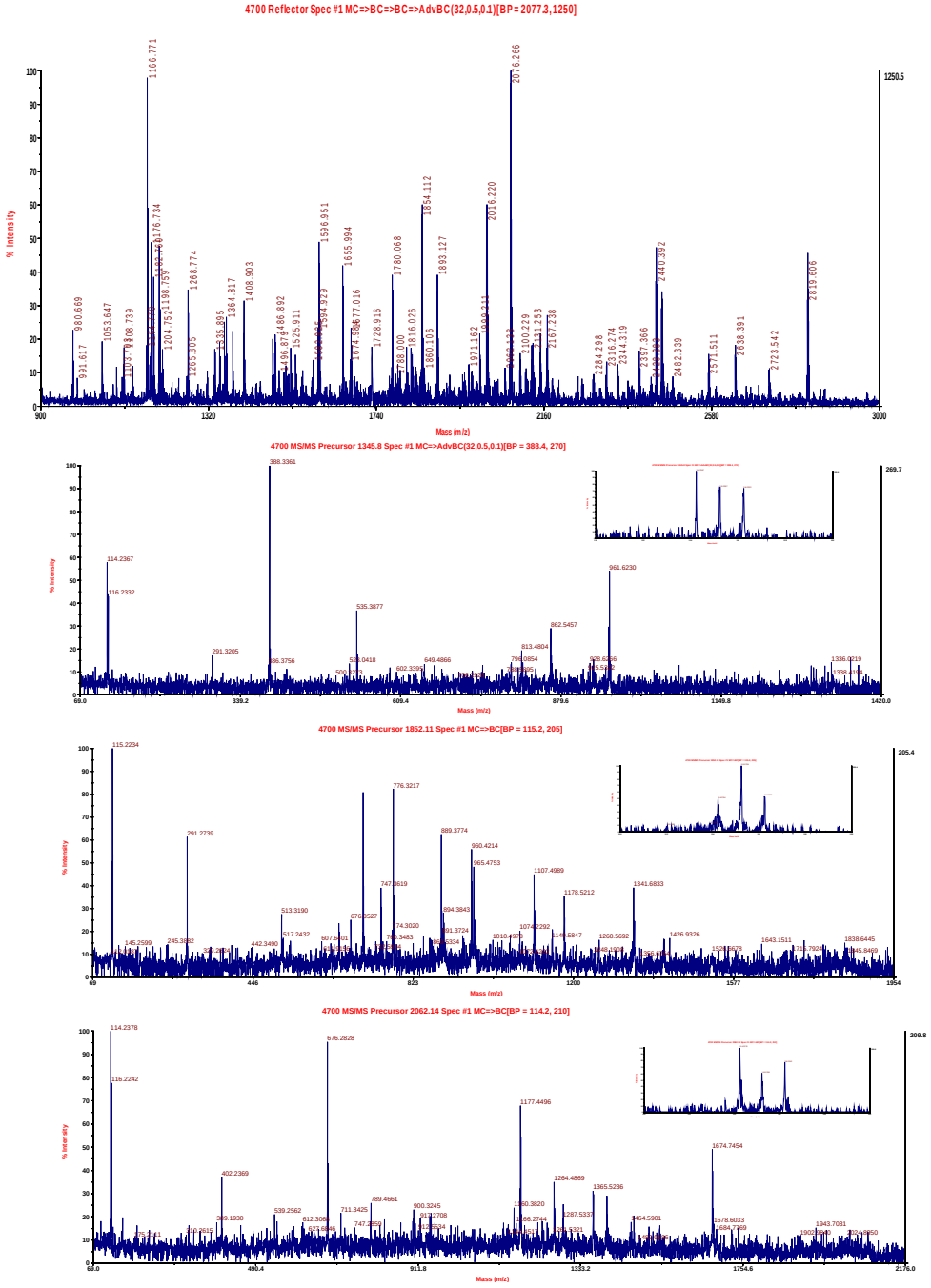


Fig 4. Examples of selected MS showing parent ions (a) and MS/MS fragment ions (b, c, and d).

Repeatability

Discussion

1. method comparison with published paper. LC/MS vs. LC/MALDI, filtration concentration vs C4 column concentration, proteins recovery with different depletion columns, fraction collection of SCX samples.

Strong cation exchange is used instead of C4 HPLC chromatography since the latter has a lower peptide yield (70% according to our analysis). The reason of this is not clear.

Repeatability

References:

1. Bunai K, Nozaki M, Hamano M, Ogane S, Inoue T, Nemoto T, Nakanishi H, Yamane K. Proteomic analysis of acrylamide gel separated proteins immobilized on polyvinylidene difluoride membranes following proteolytic digestion in the presence of 80% acetonitrile. Proteomics 2003;3:1738-1749 [PMID: 12973734].

2. Bluff JE, Brown NJ, Reed MW, Staton CA. Tissue factor, angiogenesis and tumour progression. Breast Cancer Res. 2008;10:204 [PMID: 18373885]

3. Yaffe MJ. Mammographic density. Measurement of mammographic density. Breast Cancer Res. 2008;10:209 [PMID: 18598375]

4. Hojilla CV, Wood GA, Khokha R. Inflammation and breast cancer: metalloproteinases as common effectors of inflammation and extracellular matrix breakdown in breast cancer. Breast Cancer Res. 2008;10:205

5. Isserlin R, Emili A. Interpretation of large-scale quantitative shotgun proteomic profiles for biomarker discovery. Curr Opin Mol Ther. 2008;10:231-42. [PMID: 18535930]

6. Nesvizhskii AI, Aebersold R. Interpretation of shotgun proteomic data: the protein inference problem. Mol Cell Proteomics. 2005;4:1419-40. [PMID: 16009968]

7. Garbis SD, Tyritzis SI, Roumeliotis T, Zerefos P, Giannopoulou EG, Vlahou A,
Kossida S, Diaz J, Vourekas S, Tamvakopoulos C, Pavlakis K, Sanoudou D,
Constantinides CA. Search for Potential Markers for Prostate Cancer Diagnosis, Prognosis and Treatment in Clinical Tissue Specimens Using Amine-Specific Isobaric Tagging (iTRAQ) with Two-Dimensional Liquid Chromatography and Tandem Mass Spectrometry. J Proteome Res. 2008;. [PMID: 18553995]

8. Glen A, Gan CS, Hamdy FC, Eaton CL, Cross SS, Catto JW, Wright PC, Rehman I.
iTRAQ-facilitated proteomic analysis of human prostate cancer cells identifies proteins associated with progression. J Proteome Res. 2008;7:897-907.[PMID: 18232632].

9. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol. 1999;17:994-999.[PMID: 10504701]

10. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol. 2001;19:242-247 [PMID: 11231557]

11. Liu T, Qian WJ, Mottaz HM, Gritsenko MA, Norbeck AD, Moore RJ, Purvine SO, Camp DG 2nd, Smith RD. Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. Mol Cell Proteomics. 2006;5:2167-2174.[PMID: 16854842]

12. Zolotarjova N, Mrozinski P, Chen H, Martosella J. Combination of affinity depletion of abundant proteins and reversed-phase fractionation in proteomic analysis of human plasma/serum. J Chromatogr A. 2008;1189:332-338. [PMID: 18154976]

13. Varnum SM, Covington CC, Woodbury RL, Petritis K, Kangas LJ, Abdullah MS, Pounds JG, Smith RD, Zangar RC. Proteomic characterization of nipple aspirate fluid: identification of potential biomarkers of breast cancer. Breast Cancer Res Treat. 2003;80:87-97.[PMID: 12889602]

14. Omenn GS, States DJ, Adamski M, Blackwell TW, Menon R, Hermjakob H, Apweiler R, Haab BB, Simpson RJ, Eddes JS, Kapp EA, Moritz RL, Chan DW, Rai AJ, Admon A, Aebersold R, Eng

J, Hancock WS, Hefta SA, Meyer H, Paik YK, Yoo JS, Ping P, Pounds J, Adkins J, Qian X, Wang R, Wasinger V, Wu CY, Zhao X, Zeng R, Archakov A, Tsugita A, Beer I, Pandey A, Pisano M, Andrews P, Tammen H, Speicher DW, Hanash SM. Overview of the HUPO Plasma Proteome Project: results from the pilot phase with35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. Proteomics. 2005;5:3226-3245.[PMID: 18369884]

15. Omenn GS The Human Proteome Organization Plasma Proteome Project pilot phase: Reference specimens, technology platform comparisons, and standardized data submissions and analyses Proteomics 2004, 4:1235–1240 [PMID: 16104055]